Sociedad de Estadística e Investigación Operativa
Top (2004) Vol. 12, No. 2, pp. 209–299

On Self-Regular IPMs

Maziar Salahi
Department of Mathematics and Statistics
McMaster University, Hamilton, ON, L8S 4K1, Canada
E-mail: msalahi@optlab.mcmaster.ca
Renata Sotirov, and Tamás Terlaky
Advanced Optimization Laboratory, Department of Computing & Software,
McMaster University, Hamilton, ON, L8S 4K1, Canada
E-mail: rsotirov@optlab.mcmaster.ca, terlaky@mcmaster.ca

Abstract

Primal-dual interior-point methods (IPMs) have shown their power in solving large classes of optimization problems. However, at present there is still a gap between the practical behavior of these algorithms and their theoretical worst-case complexity results, with respect to the strategies of updating the duality gap parameter in the algorithm. The so-called small-update IPMs enjoy the best known theoretical worst-case iteration bound, but work very poorly in practice. To the contrary, the so-called large-update IPMs have superior practical performance but with relatively weaker theoretical results. In this paper we discuss the new algorithmic variants and improved complexity results with respect to the new family of Self-Regular proximity based IPMs for Linear Optimization problems, and their generalizations to Conic and Semidefinite Optimization

Key Words: Linear optimization, semidefinite optimization, conic optimization, primal-dual interior-point method, self-regular proximity function, polynomial complexity.

AMS subject classification: .

1 Introduction

In 1984, Karmarkar (Karmarkar (1984)) proposed a Linear Optimization (LO) algorithm with a polynomial complexity that was able to solve large–scale LO problems more efficiently than the Simplex method. This was a breakthrough in solving large–scale LO problems and the beginning of

This research was supported by the MITACS project "New IPMs and Software for Convex Conic–Linear Optimization and Their Application to Solve VLSI Circuit Layout Problems ", by an NSERC discovery grant, and the CRC program. The first author would also like to thank the Iranian Ministry of Science, Research and Technology for supporting his research.

the era of modern interior-point methods (IPMs). In 1994, Nesterov and Nemirovski (Nesterov and Nemirovski (1994)) invented the theory of self–concordant functions, allowing polynomial time IPMs to be extended to more complex problems such as Semidefinite Optimization (SDO) and Second-Order Conic Optimization (SOCO). Nowadays IPMs offer the most efficient algorithms for SDO of reasonable size. For a survey on IPMs see the recent books (Renegar (2001), Roos et al. (1997), Wright (1997) and Ye (1997)).

In this paper, we give an overview of primal-dual IPMs based on the so–called *Self–Regular (SR) proximity functions* (Peng et al. (2002b)). We explain the SR approach for LO problems briefly, and only the basic concept for SDO and SOCO problems. The extensions of the IPMs based on SR-proximities for linear and nonlinear $P_*(\kappa)$ complementarity problems (CP) are presented in Peng et al. (2002b). This generalization is not addressed here since we limit the scope of this paper to Conic Linear Optimization (CLO).

We consider LO problems in the standard form

$$\min\{c^T x : Ax = b, x \geq 0\}, \tag{P}$$

where $A \in I\!R^{m\times n}$ satisfies $\text{rank}(A) = m$, $b \in I\!R^m$, $c, x \in I\!R^n$. The dual problem of (P) is

$$\max\{b^T y : A^T y + s = c, s \geq 0\}. \tag{D}$$

We may assume without loss of generality (see Klerk (2002) and Roos et al. (1997)) that there exists an (x^0, y^0, s^0) such that

$$Ax^0 = b,\ x^0 > 0, \qquad A^T y^0 + s^0 = c,\ s^0 > 0,$$

i.e., (P) and (D) satisfy the *interior point condition* (IPC). If the IPC holds, then finding an optimal solution of (P) and (D) is equivalent (see Wright (1997)) to solving the system of *optimality conditions*

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= 0, \end{aligned} \tag{1.1}$$

where xs denotes the coordinatewise product of the vectors x and s. The last equation in (1.1) is known as the *complementarity condition*. A basic

step toward primal-dual IPMs is to perturb the complementarity condition with some $\mu > 0$. This leads to the following system

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= \mu e, \end{aligned} \tag{1.2}$$

where $e = (1, \ldots, 1)^T$. If the IPC holds, then for each $\mu > 0$, system (1.2) has a *unique solution* $(x(\mu), y(\mu), s(\mu))$, see Güler 1994), Kojima (1989) and Roos et al. (1997). The set

$$\big\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\big\},$$

defines a smooth analytic curve parameterized by μ, called *the primal–dual central path* (Roos et al. (1997)). The limit of the central path as μ goes to zero exists, see Güler and Ye (1993), Megiddo (1986) and Sonnevend (1986). Because the limit point satisfies the complementarity condition, it naturally yields optimal solutions for both (P) and (D). Primal-dual IPMs follow the central path approximately, and generate points in a certain neighborhood of the central path. A neighborhood of the central path can be defined in the following way

$$\begin{aligned} \mathcal{N}(n, \tau) := \big\{(x, y, s) : (x, s) > 0, \ Ax = b, A^T y + s = c, \\ \Phi(x, s, \mu) \leq \eta(n, \tau), \ \mu > 0\big\}, \end{aligned} \tag{1.3}$$

where $\Phi(x, s, \mu)$ is a *proximity function* to measure the distance from the present point to the central path, and $\eta(n, \tau)$ is a function of a parameter $\tau > 0$ and the dimension n of the underlying problem. For different choices of the proximity measure for classical IPMs see Roos et al. (1997) and Wright (1997).

Now we describe how classical primal–dual IPMs work (see also Algorithm 1). We start with a point (x, y, s) that satisfies the IPC. Without loss of generality (Roos et al. (1997)) we may assume that $(x, y, s) \in \mathcal{N}(n, \tau)$ on the central path with $\mu = 1$. The aim is to move this triple toward the optimal set. In order to do that we target a point on the central path given by the target parameter $\mu := (1 - \theta)\mu$, for some $\theta \in [0, 1]$. Then we solve

the Newton system

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu e - xs, \end{aligned} \tag{1.4}$$

to obtain the search direction. Since matrix A has a full row rank, the Newton system (1.4) has a unique solution for all $\mu > 0$. Then we take a step along the search direction with a step size $\alpha \in (0, 1]$ which is defined by some line search rule. The search direction and line search rule ensure that the new triple $(x + \alpha\Delta x, y + \alpha\Delta y, s + \alpha\Delta s)$ is closer to the μ–center $(x(\mu), y(\mu), s(\mu))$. This step is repeated as long as the actual iterate is sufficiently close to the μ center. Then μ is reduced again by the factor $(1 - \theta)$ and the process is repeated until an approximate solution to the problem is obtained, e.g., until μ gets small enough.

Note that our primary goal is to reduce the *duality gap* as fast as possible. This is done by subsequently decreasing the parameter μ with a fixed ratio $1 - \theta$ at each iteration of the algorithm. As a consequence, the choice of the parameter θ has an important role in the design and analysis of IPMs. If θ is a constant, for instance $\theta = 1/2$, then we call the algorithm a *large-update* (or long-step) method. If θ depends on the problem dimension, such as $\theta = 1/\sqrt{n}$, then the algorithm is named a *small-update* (or short-step) method. In the classical primal–dual IPMs there is a gap between the practical performance of IPMs and their theoretical worst-case complexity results with respect to different choices of θ. The small-update method has the best known $\mathcal{O}(\sqrt{n}\log\frac{n}{\epsilon})$ iteration bound, while the large-update method has a worse $\mathcal{O}(n\log\frac{n}{\epsilon})$ iteration bound (Roos et al. (1997), Wright (1997) and Ye (1997)). However, large-update IPMs perform much better in practice than small-update methods (Andersen et al. (1996)).

Several strategies have been proposed to decrease the gap, i.e., to improve the complexity of large-update IPMs. Hung and Ye (1996), Jansen et al. (1997) and Monteiro et al. (1990) use higher order methods to reduce the complexity of large-update IPMs. However, there is a price to pay for the reduced complexity; higher order methods are computationally more expensive per iteration than first order methods, since some additional equation systems need to be solved with the same coefficient matrix at each iteration. Recently Peng et al. (Peng et al. (2002b) and Peng et al.

Algorithm 1

Input:
a proximity parameter τ;
an accuracy parameter $\epsilon > 0$;
(x^0, s^0) and $\mu^0 = 1$ such that $\Phi(x^0, s^0, \mu^0) \leq \eta(n, \tau)$;
begin
$x := x^0$; $s := s^0$; $\mu := \mu^0$;
while $n\mu \geq \epsilon$ **do**,
choose[a] μ;
while $\Phi(x, s, \mu) \geq \eta(n, \tau)$ **do**,
begin
solve the system (1.8) for $\Delta x, \Delta y, \Delta s$;
determine a step size α;
$x := x + \alpha\Delta x$;
$s := s + \alpha\Delta s$;
$y := y + \alpha\Delta y$,
end
end
end

[a] There are different ways to choose μ. In classical IPMs μ is decreased by the factor $1 - \theta$ at each iteration, where θ may depend on n or may be independent of n, see Roos et al. (1997). In our adaptive algorithm, μ is chosen as it is described in Section 3 (see also Salahi and Terlaky (2004a)).

(2002c)) have proposed a new strategy for improving the theoretical complexity of large-update IPMs. The choice of the proximity measure and search direction turned out to be the crucial factor for the performance, as well as for the quality of the analysis of IPMs. In this paper, we focus on the novel approach proposed in Peng et al. (2002b) and Peng et al. (2002c), where new primal–dual IPMs are induced based on so-called *Self-Regular proximity measures* that are used to define new search directions and to control the iterative process. We discuss several algorithm variants for large-update path–following methods and report their polynomial iteration complexity.

To describe the new family of algorithms, we introduce the following notation. Let $I\!R^n_{++} = \{x \in I\!R^n | x > 0\}$ denote the positive orthant and $I\!R^n_+$

denote the nonnegative orthant in $\mathbb{R}^n$. For any strictly feasible primal-dual pair (x, s) and any $\mu > 0$, we define the vectors

$$v := \sqrt{\frac{xs}{\mu}}, \quad \text{and} \quad v^{-1} := \sqrt{\frac{\mu e}{xs}}, \tag{1.5}$$

whose i^{th} components are $\sqrt{\frac{x_i s_i}{\mu}}$ and $\sqrt{\frac{\mu}{x_i s_i}}$, respectively. Using the above notation, one can state the *centrality condition* in (1.2) as

$$v = v^{-1} = e,$$

and rewrite the last equation in (1.4) as

$$d_x + d_s = v^{-1} - v, \tag{1.6}$$

where

$$d_x := \frac{v\Delta x}{x}, \quad \text{and} \quad d_s := \frac{v\Delta s}{s}. \tag{1.7}$$

The vectors d_x, d_s give the search directions in the *scaled v–space* and system (1.4) in the scaled space is given as

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v, \end{aligned} \tag{1.8}$$

where

$$\bar{A} = \frac{1}{\mu} AV^{-1}X, \tag{1.9}$$

for $V = \text{diag}(v)$ and $X = \text{diag}(x)$.

Remark 1.1. The right hand side of the third equation in (1.8) is the negative gradient of the primal-dual logarithmic barrier function that is defined as follows

$$\Psi_{\ell b}(v) = \sum_{i=1}^{n} \frac{v_i^2 - 1}{2} - \log v_i.$$

This shows that the scaled version of the classical search direction in primal–dual methods for LO can be interpreted as the steepest descent direction

for the scaled logarithmic barrier function.

The approach described in this paper is based on the observation that in principle any twice continuously differentiable strictly convex function $\psi(t)$ which satisfies the following conditions

$$\psi'(1) = \psi(1) = 0, \tag{1.10}$$

$$\psi''(t) > 0, \quad \forall\, t > 0, \tag{1.11}$$

$$\lim_{t\downarrow 0} \psi(t) = \lim_{t\to\infty} \psi(t) = \infty, \tag{1.12}$$

is a good candidate to replace the logarithmic barrier function ($\frac{t^2-1}{2} - \log t$) in defining the proximity measure and the search direction in primal-dual IPMs. Then a *proximity measure* $\Psi(v) : I\!R^n_{++} \to I\!R_+$, based on the *kernel function* ψ, in the scaled space can be defined as the sum of the componentwise deviations

$$\Psi(v) := \sum_{i=1}^{n} \psi(v_i). \tag{1.13}$$

The motivation of the new search direction is to increase the small components $v_i < 1$ and to decrease the large components $v_i > 1$, *more* than the classical one. It is reasonable to expect that such approach leads to approaching the μ–center faster. Hence, the crucial step in the new approach is to define appropriate functions that measure the discrepancy between the vectors $e = (1, \ldots, 1)^T$ and v and satisfies properties (1.10)–(1.12). Then (1.13) can be used to define a new *proximity function* and the negative gradient of $\Psi(v)$ can be used to define a new (scaled) search direction, i.e., the third equation in (1.8) may be written as

$$d_x + d_s = -\nabla\Psi(v). \tag{1.14}$$

In Peng et al. (2002b) and Peng et al. (2002c), Peng et al. introduce the class of SR functions which provides a rich source of kernel functions and satisfies all desired properties. They modify system (1.8) according to (1.14), i.e., the solution of the modified system gives the projected steepest descent direction for the proximity measure $\Psi(v)$. By this modification of the third equation in (1.8), the following Newton system is obtained:

$$\bar{A}d_x \;\; = \;\; 0,$$

$$\begin{aligned} \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= -\nabla \Psi(v). \end{aligned} \tag{1.15}$$

System (1.15) can be equivalently presented in terms of the search directions in the original space in the following way

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= -\mu v \nabla \Psi(v). \end{aligned} \tag{1.16}$$

By using a SR function as a kernel function, Peng et al. (Peng et al. (2002b) and Peng et al. (2002c)) show that the worst case iteration complexity of large-update methods can be improved so that the gap between the theoretical results for small- and large-update IPMs is significantly reduced. This surprising result follows from a careful analysis of the new method, which depends on a number of technical results. Here we present some of the main results, and for details we refer to the papers Peng et al. (2002b), Peng et al. (2002c), Peng et al. (2002d), Salahi and Terlaky (2004a) and Salahi and Terlaky (2004b).

The paper is organized as follows. In Section 2 we define univariate SR functions and give their basic properties. We also give the properties of SR-proximity functions for LO. In Section 3 we present an adaptive large-update SR–IPM for LO, and in Section 4 a SR-proximity based predictor-corrector IPM for LO. In Section 5 we explain the concept of an infeasible SR-IPM. Some closely related kernel functions that are not SR, but still allow to prove polynomial complexity of IPMs, are described in Section 6. In Sections 7 and 8, we explain how the concept of SR proximity based IPMs for LO can be extended to more complex problems, such as SDO and SOCO, respectively.

Notation. Throughout the paper $\|\cdot\|$ denotes the 2-norm of vectors and $\|\cdot\|_\infty$ denotes the infinity norm. We denote by $\mathcal{I}$ the index set $\mathcal{I} = \{1, 2, \ldots, n\}$ and $x^{-T}s^{-1} = \sum_{i\in\mathcal{I}} x_i^{-1}s_i^{-1}$. For any $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$, $x_{\min} = \min\{x_1, x_2, \ldots, x_n\}$ and $x_{\max} = \max\{x_1, x_2, \ldots, x_n\}$ denote the smallest and the largest components of x, respectively.

2 Self–Regular Proximity Functions

2.1 Univariate SR Functions

Here we give the definition and the basic properties of a univariate SR kernel functions (Peng et al. (2002b) and Peng et al. (2002c)). In Sections 7 and 8 we show how the concept of SR functions can be transparently extended to various cones, such as the cone of positive semidefinite matrices and the second–order cone, respectively.

As mentioned before, the centrality condition in the v–space (i.e., $v = e$) motivates us to construct functions in $I\!R^n_+$ that attains the global minimum at e and can be used to measure the distance from any point in $I\!R^n_+$ to e. However, it is also desirable for the function to enjoy certain barrier properties that prevents the argument from moving to the boundary of $I\!R^n_+$. Motivated with the previously described function–properties Peng et al. (Peng et al. (2002c)) define the family of univariate SR functions in the following way.

Definition 2.1. A function $\psi : (0, \infty) \to I\!R$, $\psi \in \mathcal{C}^2$ is *Self–Regular* if it satisfies the following conditions:

SR.1 $\psi(t)$ is strictly convex with respect to $t > 0$ and vanishes at its global minimal point $t = 1$, i.e., $\psi(1) = \psi'(1) = 0$. Further, there exist positive constants $\nu_2 \geq \nu_1 > 0$ and $p \geq 1,\ q \geq 1$ such that

$$\nu_1(t^{p-1} + t^{-1-q}) \leq \psi''(t) \leq \nu_2(t^{p-1} + t^{-1-q}), \quad \forall t \in (0, \infty); \quad (2.1)$$

SR.2 For any $t_1, t_2 > 0$,

$$\psi(t_1^r t_2^{1-r}) \leq r\psi(t_1) + (1 - r)\psi(t_2), \quad \forall r \in [0, 1]. \quad (2.2)$$

If $\psi(t)$ is SR, then parameter q is called the *barrier degree*, and parameter p is called the *growth degree* of the SR function $\psi(t)$. There are two popular families of SR functions that are used in the design of new SR-IPMs and play a crucial role in the analysis of the new search directions.

Example 2.1. The first family is given by

$$\Upsilon_{p,q}(t) = \frac{t^{p+1} - 1}{p(p + 1)} + \frac{t^{1-q} - 1}{q(q - 1)} + \frac{p - q}{pq}(t - 1), \quad p \geq 1,\ q > 1, \quad (2.3)$$

with $\nu_1 = \nu_2 = 1$. The second family is defined as

$$\Gamma_{p,q}(t) = \frac{t^{p+1}-1}{p+1} + \frac{t^{1-q}-1}{q-1}, \quad p \geq 1,\ q > 1, \tag{2.4}$$

with $\nu_1 = \min(p,q)$ and $\nu_2 = \max(p,q)$.

Equivalent conditions to condition SR.2 are given in the following lemma.

Lemma 2.1 (Peng et al. (2002b)). *Let $\psi : (0,\infty) \to [0,\infty), \psi \in \mathcal{C}^2$. Function $\psi(\cdot)$ satisfies SR.2 if and only if the following equivalent statements hold.*

1) $\psi(e^t)$ is convex.

2) $\psi^{'}(t) + t\psi^{''}(t) \geq 0$ when $t > 0$.

3) $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}\psi(t_1) + \frac{1}{2}\psi(t_2)$, for $t_1, t_2 > 0$.

The following property of SR functions can be easily proved.

Proposition 2.1 (Peng et al. (2002b)). *If the functions $\psi_1(t), \psi_2(t)$ are SR, then so is any convex conic combination $\beta_1\psi_1 + \beta_2\psi_2$ with $\beta_1, \beta_2 \geq 0,\ \beta_1 + \beta_2 > 0$.*

Since any nontrivial conic combination of two SR functions ψ_1 and ψ_2 is SR, the set of SR functions is a pointed convex cone. The following proposition collects several properties of the SR functions that are proved in Peng et al. (2002b).

Proposition 2.2. *Let Ω_1, Ω_2 respectively, be the sets of functions whose elements satisfy conditions SR.1 and SR.2, respectively. For $t > 0$ the following statements hold.*

1. *Suppose that $\psi(t) \in \Omega_1$. Then $\left|\frac{1}{t}\psi'(t)\right| \leq \frac{\nu_2}{\nu_1}\psi''(t)$.*
2. *If a function $\psi(t) \in \Omega_1$ with $\nu_1 = \nu_2$, then it is SR.*
3. *Suppose that $\psi(t) \in \Omega_1$. Then $2\nu_1\psi(t) \leq \psi(t)'(t)^2$.*
4. *If $\psi(t) = \psi(t^{-1})$ and $\psi(t) \in \Omega_1$, then $\psi(t)$ is SR.*
5. *If $\psi(t) \in \Omega_2$, then also $\psi\left(\frac{1}{t}\right) \in \Omega_2$.*
6. *Let N be any positive integer and $\psi(t) = \beta_0 \log t + \sum_{i=1}^{N} \beta_i(t^{\rho_i} - 1), \quad \beta_0 \in \mathbb{R},\ \beta_i \geq 0,\ \rho_i \in \mathbb{R},\ i = 1, 2, \ldots, N$. Then $\psi(t) \in \Omega_2$.*

2.2 Properties of SR-Proximity Functions for LO

Here we present some properties of SR-proximity functions that will be used in the analysis of SR-IPMs. First we study common properties of the proximity functions that are based on the $\Gamma_{1q}(t)$ family for $q \in I\!R,\ q > 1$, and then separately the properties that apply only for $q = 3$ and $q = \log n + 1$. SR-proximity functions induced by the $\Gamma_{1q}(t),\ q \in I\!R,\ q > 1$ family of kernel functions allow to design various large-update algorithms, with the best known iteration complexity (Peng and Terlaky (2002), Salahi et al. (2003) and Salahi and Terlaky (2004a)).

Let us define the family of proximity functions based on the $\Gamma_{1q}(t)$ family as

$$\Phi_q(x, s, \mu) := \Psi_q(v) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{1-q} - n}{q - 1}. \tag{2.5}$$

Proposition 2.3 (Salahi and Terlaky (2004a)). *For any fixed $(x, s) > 0$, the proximity function $\Phi_q(x, s, \mu)$ as a function of μ, attains its global minimum at*

$$\mu_q^* = \left(\frac{x^T s}{(x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}}} \right)^{\frac{2}{q+1}}.$$

The following result is an immediate consequence of Proposition 2.3.

Corollary 2.1 (Salahi and Terlaky (2004a)). *For any fixed $(x, s) > 0$, the proximity function $\Phi_q(x, s, \mu)$ is a decreasing function with respect to μ when $\mu \leq \mu_q^*$, and it is an increasing function of μ if $\mu > \mu_q^*$.*

The following lemma plays an important role in the definition of the SR neighborhood.

Lemma 2.2 (Salahi and Terlaky (2004a)). *Let $\tau \in I\!R,\ \tau \geq 2$. For the generalized harmonic mean*

$$\mu_q^h = \left(\frac{n}{(x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}}} \right)^{\frac{2}{q-1}}$$

of the components of vector xs and the current duality gap

$$\mu_g := \frac{x^T s}{n}, \tag{2.6}$$

the following statements are equivalent:

1) $\frac{\mu_g}{\mu_q^h} \leq \tau$,

2) $\Phi_q\left(x, s, \frac{\mu_g}{\tau}\right) \leq \frac{(\tau-1)n}{2}$,

3) $\Phi_q(x, s, \mu_g) \leq \frac{\left(\tau^{\frac{q-1}{2}}-1\right)n}{q-1}$.

Due to the choice of the kernel function $\psi(t) = \Gamma_{1q}(t)$, system (1.16) has the following form

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= \mu^{\frac{q+1}{2}}(x^{\frac{1-q}{2}})s^{\frac{1-q}{2}} - xs. \end{aligned} \tag{2.7}$$

Let us denote the solution of system (2.7) by $\big(\Delta x(\mu), \Delta y(\mu), \Delta s(\mu)\big)$. The following two lemmas discuss the change of the duality gap along the search direction $\big(\Delta x(\mu), \Delta y(\mu), \Delta s(\mu)\big)$ for $\mu = \mu_q^*$ and $\mu = \mu_q^h$.

Lemma 2.3 (Salahi and Terlaky (2004a)). *Let* $\big(\Delta x(\mu_q^*), \Delta y(\mu_q^*), \Delta s(\mu_q^*)\big)$ *be the solution of system (2.7) with* $\mu = \mu_q^*$. *Then the relation*

$$x^T\Delta s(\mu_q^*) + s^T\Delta x(\mu_q^*) = 0$$

holds.

Proof. Directly from the definition of μ_q^*. □

Corollary 2.2 (Salahi and Terlaky (2004a)). *If* $\mu = \mu_q^*$, *then the duality gap will not change for any feasible step size* α, *i.e.,*

$$(x + \alpha\Delta x(\mu_q^*))^T(s + \alpha\Delta s(\mu_q^*)) = x^Ts.$$

Analogous to Lemma 2.3, the following result is obtained for $\mu = \mu_q^h$.

Lemma 2.4 (Salahi and Terlaky (2004a)). *Let* $\big(\Delta x(\mu_q^h), \Delta y(\mu_q^h), \Delta s(\mu_q^h)\big)$ *be the solution of system (2.7) with* $\mu = \mu_q^h$. *Then the relation*

$$x^T\Delta s(\mu_q^h) + s^T\Delta x(\mu_q^h) = n\mu_q^h - x^Ts$$

holds.

Corollary 2.3. *If the targeted parameter is μ_q^h, then the search direction based on our specific SR-proximity function and the standard Newton direction will predict the change of the duality gap in the same way, i.e.,*

$$(x + \alpha\Delta x(\mu_q^h))^T(s + \alpha\Delta s(\mu_q^h)) = x^T s\left(1 - \alpha + \frac{\mu_q^h \alpha}{\mu_g}\right).$$

In some of the algorithms described in the subsequent sections, we explore properties of some proximity functions based on particular members of the $\Gamma_{1q}(t)$ family i.e., for $q = 3$ and $q = \log n + 1$. Since these proximity functions have specific properties that do not apply for arbitrary $q > 1$, we present them separately.

- $q = 3$

  The SR function $\Gamma_{13}(t)$ is used as a kernel function in the infeasible algorithm (see Section 5). The proximity function based on the kernel function $\Gamma_{13}(t)$ is defined as follows

  $$\Phi_3(x, s, \mu) := \Psi_3(v) = \frac{1}{2}\|v - v^{-1}\|^2. \tag{2.8}$$

  The following proposition shows that for fixed $(x, s) > 0$, the function $\Phi_3(x, s, \mu)$ attains its global minimum at a point $\mu < \mu_g$, where μ_g is the current duality gap defined by (2.6).

  Proposition 2.4 (Peng and Terlaky (2002)). *For any fixed $(x, s) > 0$, the proximity function $\Phi_3(x, s, \mu)$ as a function of μ, has the global minimizer μ_3^* that is the geometric mean of μ_g and μ_3^h, i.e.,*

  $$\mu_3^* = \sqrt{\frac{x^T s}{x^{-T} s^{-1}}} = \sqrt{\mu_g \mu_3^h}.$$

  Note that Proposition 2.4 directly follows from Proposition 2.3 and Lemma 2.2.

  It is easy to verify the following interesting relations that play a crucial role in the design of algorithmic schemes based on this special proximity function.

Proposition 2.5 (Peng and Terlaky (2002)). *Suppose that* $(x, s) > 0$ *is fixed. Then we have*

$$\Phi(x, s, \mu_g) = \Phi(x, s, \mu_3^h),$$

and

$$\Phi(x, s, \mu_g) = \Phi(x, s, \mu_3^*) + \frac{\Phi(x, s, \mu_3^*)^2}{2n}.$$

Based on the $\Gamma_{13}(t)$ kernel function, the authors of Peng and Terlaky (2002) propose also a dynamic large-update IPM which improve the complexity of large-update methods significantly.

- $q = 1 + \log n$

  In the SR predictor–corrector approach, that is presented in Section 4, the following SR-proximity function is used

$$\Phi_\ell(x, s, \mu) := \Psi_\ell(v) = \frac{e^T v^2 - n}{2} + \frac{e^T v^{-\log n} - n}{\log n}. \tag{2.9}$$

  For fixed $(x, s) > 0$, the global minimum of the proximity function (2.9) with respect to μ is $\mu_l^* := \mu_{1+\log n}^*$ (see Proposition 2.3). We provide some specific properties of this proximity function in Section 4.

3 Adaptive Large-update IPM for LO

In this section we present a family of adaptive large-update IPMs for LO based on the $\Gamma_{1q}(t), q > 1$ family of kernel functions (Salahi and Terlaky (2004a)). The motivation for our adaptive algorithm is to develop a variant, that is more flexible in updating μ than classical polynomial IPMs (Roos et al. (1997)), and is closer to what is implemented in IPM solvers (LIPSOL, Zhang (1999), and McIPM, Zhu (2003)). The algorithm described here use large-update at each iteration and does not make any recentering iteration.

For $\tau \geq 2$, we set $\frac{(\tau-1)n}{2}$ as the maximum allowed value of the proximity function. At each iteration we choose the target value μ_q^t so that $\Phi_q(x, s, \mu_q^t) = \frac{(\tau-1)n}{2}$. One can see that $\Phi_q(x, s, \mu_q^t) = \frac{(\tau-1)n}{2}$ if and only if

μ_q^t solves the equation

$$2(x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}} \mu^{\frac{q+1}{2}} - (2n + \tau(q-1)n)\,\mu + (q-1)x^T s = 0. \tag{3.1}$$

This equation has two positive roots if and only if $\Phi_q(x, s, \mu_g) \leq \frac{(\tau^{\frac{q-1}{2}}-1)n}{q-1}$. One of them is less than or equal to μ_q^*, and the other is larger than or equal to μ_q^*. At each iteration of the algorithm, we use the smaller positive root μ_q^t as the target value. One can easily prove that $\mu_q^t \leq \mu_q^h$ holds when $\mu_g \leq \tau\mu_q^h$ with $\mu_q^t = \mu_q^h$ if and only if $\mu_g = \tau\mu_q^h$.

The following lemma gives a lower bound for the norm of $(v - v^{-q})$, when the target value is μ_q^t.

Lemma 3.1. *Let $\tau \geq 2$. Then $\sigma = \|v - v^{-q}\| \geq 1$.*

Proof. By using Proposition 3.1.5 of Peng et al. (2002b), we get

$$\sigma^2 \geq 2\Phi(x, s, \mu_q^t) = (\tau - 1)n \geq 1, \quad \forall n \geq 1.$$

□

Lemma 3.2 (Salahi and Terlaky (2004a)). *Let μ_q^t be the smaller positive root of equation (3.1). Then the inequality*

$$\tau\mu_q^t \leq \mu_g \leq (\tau + \frac{2}{q-1})\mu_q^t$$

holds.

Proof. We give a proof for the right hand side inequality. The left hand side can be proved analogously. The function in (3.1) is a convex function w.r.t. μ. If we replace μ by $\frac{\mu_g}{\tau+\frac{2}{q-1}}$ in that function and simplify, it suffices to prove that

$$2\left(\frac{\mu_g}{(\tau + \frac{2}{q-1})\mu_q^h}\right)^{\frac{q-1}{2}} \frac{1}{\tau + \frac{2}{q-1}} - \frac{2 + \tau(q-1)}{\tau + \frac{2}{q-1}} + q - 1 \geq 0, \tag{3.2}$$

which is obviously true. This completes the proof of the right hand side inequality and the proof of the lemma too. □

When μ_q^t is the target value, we have the following result.

Lemma 3.3 (Salahi and Terlaky (2004a)). *Let* $\left(\Delta x(\mu_q^t), \Delta y(\mu_q^t), \Delta s(\mu_q^t)\right)$ *be the solution of system (2.7), for* $\mu = \mu_q^t$ *where* μ_q^t *is the smallest positive root of equation (3.1). Then the relations*

$$x^T \Delta s(\mu_q^t) + s^T \Delta x(\mu_q^t) = {\mu_q^t}^{\frac{q+1}{2}} (x^{\frac{1-q}{2}})^T s^{\frac{1-q}{2}} - x^T s,$$

and

$$(x + \alpha\Delta x(\mu_q^t))^T (s + \alpha\Delta s(\mu_q^t)) = (x + \alpha\Delta x)^T (s + \alpha\Delta s) = x^T s \left(1 - \alpha + \frac{{\mu_q^t}^{\frac{q+1}{2}} \alpha}{\mu_g {\mu_q^h}^{\frac{q-1}{2}}}\right) \quad (3.3)$$

hold.

The proof of Lemma 3.3 is analogous to the proofs of Lemma 2.3 and Lemma 2.4.

Remark 3.1. If $\mu_q^t \sim \mu_q^h$, then (3.3) implies that the search direction based on our specific SR-proximity function and the standard Newton direction will predict the change of the duality gap almost in the same way. (If $\mu_q^t = \mu_q^h$, see Lemma 2.4.) But if $\mu_q^t \ll \mu_q^h$, then the ratio $\frac{{\mu_q^t}^{\frac{q+1}{2}}}{\mu_g {\mu_q^h}^{\frac{q-1}{2}}}$ is very small and for the SR search direction the duality gap reduction is much larger than it would be when using the standard Newton direction.

Further motivation is available in Salahi and Terlaky (2004a). The new adaptive algorithm (see Algorithm 2) is the special case of Algorithm 1 on page 213. Algorithm 2, regardless if the iterate is close to or far away from the central path, always makes a large-update of the central path parameter μ. We choose $\mu = \mu_q^t$ at each iteration, and we just solve one Newton system at each iteration.

Let us define the SR-proximity based neighborhood (see (1.3)) as it follows:

$$\mathcal{N}_q(n,\tau) := \{(x,y,s) : (x,s) > 0,\ Ax = b, A^T y + s = c, \Phi_q(x,s,\mu_g) \le \eta_q(n,\tau),\ \mu > 0\}, \quad (3.4)$$

Algorithm 2: An Adaptive Large-Update SR-IPM

Input:
a proximity parameter $\tau \geq 2$;
an accuracy parameter $\epsilon > 0$;
$(x, s) = (x^0, s^0)$ such that $\frac{\mu_g}{\mu_q^h} \leq \tau$;
begin
while $x^T s \geq \epsilon$ **do**
$\mu := \mu_q^t$ computed from (3.1);
solve system (2.7) for $\Delta x, \Delta y, \Delta s$;
begin
determine a step size α such that
$\Phi_q(x(\alpha), s(\alpha), \mu_q^t) \leq \Phi_q(x, s, \mu_q^t) - \frac{2^{\frac{q-1}{2q}} \Phi_q(x,s,\mu_q^t)^{\frac{q-1}{2q}}}{24q}$
and $\mu_g(\alpha) \leq \tau \mu_q^h(\alpha)$;
$x := x + \alpha \Delta x$;
$y := y + \alpha \Delta y$;
$s := s + \alpha \Delta s$,
end
end
end

where $\Phi_q(x, s, \mu)$ is defined by (2.5), $\tau \geq 2$ and we define

$$\eta_q(n, \tau) := \frac{(\tau^{\frac{q-1}{2}} - 1)n}{q - 1}.$$

Recall the results of Lemma 2.2 that quantify the relation between μ_g, μ_q^h and the corresponding proximity values. These relations provide useful tools for the complexity analysis. The key element of the analysis is to give a bound for the step size α that imply sufficient reduction of μ_g. For this, one needs to explore the changing behavior of the functions $\Phi_q(x(\alpha), s(\alpha), \mu_g(\alpha))$ and $\Phi_q(x(\alpha), s(\alpha), \mu_q^t)$.

First we prove the following technical lemma.

Lemma 3.4. *Let* $\sigma = \|\nabla \Psi_q(v)\|$. *Then*

$$v_{\min} \geq (1 + \sigma)^{-\frac{1}{q}}.$$

Proof. The lemma is trivial if $v_{\min} \geq 1$. Now consider the case when $v_{\min} < 1$. Then $\sigma = \|v - v^{-q}\| \geq v_{\min}^{-q} - v_{\min} \geq v_{\min}^{-q} - 1$. This completes the proof of the lemma. □

Now we give a lower bound for the maximal feasible step size.

Lemma 3.5 (Salahi and Terlaky (2004a)). *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (2.7) where $\mu = \mu_q^t$ is the smallest positive root of equation (3.1), and $\sigma = \|\nabla \Psi_q(v)\|$. Then the maximal feasible step size, $\alpha_{\max}$, satisfies*

$$\alpha_{\max} \geq \bar{\alpha} := \frac{1}{\sigma(1+\sigma)^{\frac{1}{q}}}.$$

Proof. By using the definition of $v(\alpha)$ (see (1.5) for the definition of v), we have

$$v(\alpha_{\max}) = (v+\alpha_{\max} d_x)^{\frac{1}{2}}(v+\alpha_{\max} d_s)^{\frac{1}{2}} = v(e+\alpha_{\max} v^{-1} d_x)^{\frac{1}{2}}(e+\alpha_{\max} v^{-1} d_s)^{\frac{1}{2}}.$$

One has $v(\alpha_{\max}) \geq 0$ if

$$e + \alpha_{\max} v^{-1} d_x \geq 0 \quad \text{and} \quad e + \alpha_{\max} v^{-1} d_s \geq 0.$$

These inequalities imply

$$\alpha_{\max} \geq \frac{1}{\|(v^{-1} d_x, v^{-1} d_s)\|}.$$

We also know that

$$\|(v^{-1} d_x, v^{-1} d_s)\| \leq \frac{\|(d_x, d_s)\|}{v_{\min}} \leq \sigma(1+\sigma)^{\frac{1}{q}},$$

where the last inequality follows from Lemma 3.4. This completes the proof of the lemma. □

The following theorem estimates the decrease of the proximity function after an iterate.

Theorem 3.1 (Salahi and Terlaky (2004a)). *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (2.7), where $\mu = \mu_q^t$ is the smaller positive root of equation (3.1). Then for step size $\alpha^* := \frac{\bar{\alpha}}{3q}$, where $\bar{\alpha}$ is defined in Lemma 3.5,*

the relation

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t) \leq \Phi_q(x, s, \mu_q^t) - \frac{2^{\frac{q-1}{2q}} \Phi_q(x, s, \mu_q^t)^{\frac{q-1}{2q}}}{24q}$$

holds.

Proof. Let

$$\begin{aligned} h(\alpha) :&= \Phi_q\left(x(\alpha), s(\alpha), \mu_q^t\right) - \Phi_q(x, s, \mu_q^t) \\ &= \frac{\|v(\alpha)\|^2 - n}{2} + \frac{\|v(\alpha)^{\frac{q-1}{2}}\|^2 - n}{q-1} - \frac{\|v\|^2 - n}{2} - \frac{\|v^{\frac{q-1}{2}}\|^2 - n}{q-1}, \end{aligned} \tag{3.5}$$

where v is defined by (1.5). Using condition SR.2 of Definition 2.1 we have

$$\begin{aligned} h(\alpha) \leq \frac{1}{2} v^T (d_x + d_s)\alpha &+ \frac{1}{2(q-1)} \sum_{i=1}^{n} (v_i + \alpha(d_x)_i)^{1-q} \\ &+ \frac{1}{2(q-1)} \sum_{i=1}^{n} (v_i + \alpha(d_s)_i)^{1-q} - \frac{1}{q-1} \|v^{\frac{1-q}{2}}\|^2 := h_1(\alpha). \end{aligned} \tag{3.6}$$

It can be shown that

$$h_1'(0) = -\frac{\sigma^2}{2} \quad \text{and} \quad h_1''(\alpha) \leq \frac{q\sigma^2}{2} (v_{\min} - \alpha\sigma)^{-1-q}.$$

Then we have

$$h(\alpha) \leq -\frac{\alpha\sigma^2}{2} + \frac{q\sigma^2}{2} \int_0^\alpha \int_0^\zeta (v_{\min} - \eta\sigma)^{-1-q} d_\eta d_\zeta := h_2(\alpha).$$

Function $h_2(\alpha)$ is twice differentiable convex function on the interval $[0, \bar{\alpha})$. Let us denote by α_1^* the global minimum of $h_2(\alpha)$ on the interval $[0, \bar{\alpha})$. Then α_1^* is the unique solution of the equation

$$-\sigma^2 + \sigma\left((v_{\min} - \alpha\sigma)^{-q} - (v_{\min})^{-q}\right) = 0.$$

By a proof analogous to the proof of Lemma 3.3.3 in Peng et al. (2002b), one can show that $\alpha_1^* \geq \alpha^* = \frac{1}{3q\sigma(1+\sigma)^{\frac{1}{q}}}$. Furthermore, from Lemma 1.3.3

in Peng et al. (2002b) and for any α^* it follows that

$$h(\alpha^*) \leq -\frac{\sigma^{\frac{q-1}{q}}}{24q}.$$

We also have $\sigma^2 \geq 2\Phi_q(x, s, \mu_q^t)$ (see Proposition 3.1.5, Peng et al. (2002b), that completes the proof. □

We proceed to estimate the proximity function $\Phi_q(x(\alpha^*), s(\alpha^*), \mu_g(\alpha^*))$ when μ_q^t is used in adaptive algorithm as the targeted parameter.

Theorem 3.2 (Salahi and Terlaky (2004a)). *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (2.7), where $\mu = \mu_q^t$ is the smaller positive root of equation (3.1). Then for α^*, the relation*

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_g(\alpha^*)) \leq \frac{\left(\tau^{\frac{q-1}{2}} - 1\right) n}{q-1}$$

holds.

Proof. By Theorem 3.1, α^* is strictly feasible and

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t) \leq \frac{(\tau - 1)n}{2}.$$

Since $\Phi_q(x(\alpha^*), s(\alpha^*), \mu)$ is a quasi-convex function of μ and it converges to infinity if μ goes to zero, then there exist a $\mu_t(\alpha^*)$ such that

$$\Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t(\alpha^*)) = \frac{(\tau - 1)n}{2},$$

that is equivalent to the statement of the theorem. □

To obtain an upper bound for the total number of iterations of the algorithm, one need to estimate the change of the parameter μ_q^t before and after an iterate. The following technical lemma is needed for the complexity analysis.

Lemma 3.6 (Salahi and Terlaky (2004a)). *Let $v_+ = \frac{v}{\sqrt{1-\theta}}$ for some*

$\theta \in (0,1)$. *Then we have:*

$$\Psi_q(v_+) \leq \frac{\Psi_q(v)}{1-\theta} + \frac{n\theta}{2(1-\theta)} + \frac{\theta n}{1-\theta}\left(\frac{1-(\frac{1}{\tau+1})^{\frac{q-1}{2}}}{q-1}\right).$$

By applying Lemma 3.6 to Theorem 3.1, the following theorem can be proved.

Theorem 3.3 (Salahi and Terlaky (2004a)). *Let $\tau \geq 2$, and let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (2.7) with $\mu = \mu_q^t$ where μ_q^t is the smallest positive root of equation (3.1), and let α^* be the default step size as defined in Theorem 3.1. Then*

$$nn\Phi_q\left(x(\alpha^*), s(\alpha^*), (1-\theta)\mu_q^t\right) \qquad \leq \Phi_q(x, s, \mu_q^t), \tag{3.7}$$

nn where

$$\theta = \frac{(\tau-1)^{\frac{q-1}{2q}}}{12q\left(\tau + \log(\tau+1)\right) n^{\frac{q+1}{2q}}}.$$

Now we proceed to present the complexity of our adaptive algorithm. By the choice of μ_q^t we know that the proximity function $\Phi_q(x, s, \mu_q^t)$ keeps invariant for all the iterates. Let us denote by $\mu_q^t(\alpha^*)$ the target parameter value after one step with step size α^*. Then we have

$$\Phi_q(x, s, \mu_q^t) = \Phi_q(x(\alpha^*), s(\alpha^*), \mu_q^t(\alpha^*)).$$

On the other hand, from Theorem 3.3 we have

$$\Phi_q(x\big(\alpha^*\big), s(\alpha^*), (1-\theta)\mu_q^t\big) \leq \Phi_q\big(x(\alpha^*), s(\alpha^*), \mu_q^t(\alpha^*)\big).$$

Since the proximity function is a convex function w.r.t. μ, we have

$$\mu_q^t(\alpha^*) \leq \left(1 - \frac{(\tau-1)^{\frac{q-1}{2q}} n^{\frac{-q-1}{2q}}}{12q\left(\tau + \log(\tau+1)\right)}\right)\mu_q^t. \tag{3.8}$$

Now we are ready to present the complexity of our adaptive algorithm.

Theorem 3.4 (Salahi and Terlaky (2004a)). *Let $\tau \geq 2$. Then after at*

most

$$\left\lceil \frac{12q\left(\tau+\log(\tau+1)\right)n^{\frac{q+1}{2q}}}{(\tau-1)^{\frac{q-1}{2q}}}\log\frac{2n\tau^2}{\epsilon}\right\rceil$$

iterations the adaptive algorithm will terminate with a feasible solution satisfying $x^Ts\leq\epsilon$.

Proof. In light of inequality (3.8) we know that after at most

$$\left\lceil \frac{12q\left(\tau+\log(\tau+1)\right)n^{\frac{q+1}{2q}}}{(\tau-1)^{\frac{q-1}{2q}}}\log\frac{n(\tau+1)}{\epsilon}\right\rceil$$

iterations we have $\mu_q^t\leq\frac{\epsilon}{n(\tau+1)}$. By using Lemma 2.2 and Lemma 3.2, we have that $\mu_g\leq(\tau+1)\mu_q^t\leq\frac{\epsilon}{n}$, or equivalently $x^Ts\leq\epsilon$. □

The following corollary gives, so far, the best complexity for large-update IPMs (Peng et al. (2002b)).

Corollary 3.1 (Salahi and Terlaky (2004a)). *For* $q=\log n$, *Theorem 3.4 provides the following upper bound for the total number of iterations:*

$$\mathcal{O}\left(\sqrt{n}\log n\log\frac{n}{\epsilon}\right).$$

4 SR-Proximity Based Predictor-Corrector IPM (SR-PC)

In this section we present a SR-proximity based predictor-corrector IPM for LO (Salahi and Terlaky (2004b)). The proximity function that we use in this section is the specific proximity function $\Phi_\ell(x,s,\mu)$ introduced by (2.9). Note that for this case $q=1+\log n$.

The particular interest of this algorithm is when the iterates are far away from the central path. Predictor-corrector algorithms traditionally use an infinity neighborhood, which is defined by

$$\mathcal{N}_\infty^-(\rho):=\{(x,y,s):(x,s)>0,\ Ax=b,\ A^Ty+s=c,\\ \|(v^2-e)^-\|_\infty\leq\rho\},\quad(4.1)$$

where $a^- = \min(a, 0)$, and $\rho \in (0, 1)$ is a constant independent of n and τ. We use SR-neighborhoods, and in order to get comparable results we change $\eta(n, \tau)$ in the definition (3.4) of the neighborhood $\mathcal{N}_q(n, \tau)$ so that the resulting SR-neighborhood contains the $\mathcal{N}_\infty^-(\rho)$ neighborhood (Peng et al. (2003)).

Let $(x, s) \in \mathcal{N}_\infty^-(\rho)$, and $\mu = \mu_g$. Then for $\Phi_\ell(x, s, \mu)$ as introduced by (2.9) the following inequality is satisfied,

$$\Phi_\ell(x, s, \mu_g) = \frac{e^T v^{-\log n} - n}{\log n} \leq \frac{n(1-\rho)^{\frac{-\log n}{2}} - n}{\log n} = \frac{n^\tau - n}{\log n} = \eta_\ell(n, \tau),$$

where $\tau = 1 - \frac{1}{2}\log(1-\rho)$. With this choice of $\eta_\ell(n, \tau)$ the neighborhood $\mathcal{N}_\ell(n, \tau) := \mathcal{N}_{1+\log n}(n, \tau)$ contains the neighborhood $\mathcal{N}_\infty^-(\rho)$, and the inclusion $\mathcal{N}_\infty^-(\rho) \subseteq \mathcal{N}_\ell(n, \frac{1}{1-\rho})$ holds too. This demonstrates that these two neighborhood almost match each other.

The following remark is crucial for the rest of this section.

Remark 4.1. Without loss of generality we may always assume that after a predictor step the new iteration is on the boundary of the SR neighborhood, or equivalently $\Phi(x, s, \mu_g) = \eta_\ell(n, \tau)$.

The following result specifies the relation between μ_g and μ_ℓ^*.

Lemma 4.1 (Salahi and Terlaky (2004b)). *Let $\mu_\ell^* = \mu_{\log n+1}^*$, as defined in Proposition 2.3. If $\Phi_\ell(x, s, \mu_g) = \eta_\ell(n, \tau)$, then $\mu_\ell^* = \theta_1(\tau)\mu_g$, where $\theta_1(\tau) = n^{\frac{2(1-\tau)}{\log n+2}}$.*

Proof. Let $\mu_\ell^* = \theta_1\mu_g$ and we want to derive a lower bound for θ_1. The global minimum of the proximity function is the solution of the following equation

$$-\frac{x^T s}{2(\mu_\ell^*)^2} + \frac{(\mu_\ell^*)^{\frac{\log n-2}{2}} \left(x^{-\frac{\log n}{2}}\right)^T s^{-\frac{\log n}{2}}}{2} = 0.$$

This is equivalent to

$$\frac{n}{{\theta_1}^2\mu_g} = {\theta_1}^{\frac{\log n-2}{2}} {\mu_g}^{\frac{\log n-2}{2}} \left(x^{-\frac{\log n}{2}}\right)^T s^{-\frac{\log n}{2}},$$

and that implies

$$n{\theta_1}^{\frac{-\log n-2}{2}} = \|v^{\frac{-\log n}{2}}\|^2 = n^\tau.$$

Therefore,

$$\theta_1 = n^{\frac{2(1-\tau)}{\log n+2}}.$$

Finally, by letting $\theta_1(\tau) = \theta_1$ the proof is completed. □

In the original predictor-corrector algorithm (Mizuno et al. (1993)), *the primal-dual affine scaling direction* is used in the predictor step. The so-called primal-dual affine scaling search direction is the solution of the system

$$\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ s\Delta x + x\Delta s &= -xs. \end{aligned} \tag{4.2}$$

It is well known that close to optimality, the step size of the affine scaling direction is converging to one. This fact implies that close to optimality the affine scaling direction is the best choice.

In this section we present an SR-PC algorithm that in the first iterations chooses SR directions not only in the corrector, but also in the predictor step (Salahi and Terlaky (2004b)). The duality gap reduction of the affine scaling direction is comparable with the duality gap reduction of the SR direction, and the new iterate moves in the direction that gives larger duality gap reduction. Our next task is to define the new target value that the SR-PC algorithm use in the predictor step (when the predictor direction is a SR-direction).

Since the SR-PC algorithm operates in a large neighborhood, the maximum allowed value of the proximity function w.r.t. the target μ is $\frac{(\tau_1-1)n}{2}$, where $\tau_1 = (1-\rho)^{-1}$, see Salahi and Terlaky (2004b). Analogous to the previous section, $\Phi_\ell(x, s, \mu_\ell^t) = \frac{(\tau_1-1)n}{2}$ if and only if μ_ℓ^t is a root of

$$2(x^{\frac{-\log n}{2}})^T s^{\frac{-\log n}{2}} \mu^{\frac{\log n+2}{2}} - (2n + \tau_1 n \log n)\,\mu + x^T s \log n = 0. \tag{4.3}$$

This equation has two positive roots if and only if $\Phi_\ell(x, s, \mu_g) \leq \eta_\ell(n, \tau)$. One of the roots is less than or equal to μ_g, and the other one is larger than or equal to μ_g. In the SR-PC algorithm the smaller positive root μ_ℓ^t is used as the target value in the predictor step when the SR direction is the predictor direction. One can easily prove that $\mu_\ell^t \leq \mu_\ell^h$, holds for $\mu_g \leq \tau_1 \mu_\ell^h$

with $\mu_\ell^t = \mu_\ell^h$ if and only if $\mu_g = \tau_1 \mu_\ell^h$, where

$$\mu_\ell^h = \left(\frac{n}{\left(x^{\frac{-\log n}{2}}\right)^T s^{\frac{-\log n}{2}}} \right)^{\frac{2}{\log n}}$$

is the generalized harmonic mean of the components of the vector xs, see Salahi and Terlaky (2004a).

Although in our SR-PC algorithm the neighborhood is defined by $\Phi_l(x, s, \mu)$, we develop a family of SR-PC-IPMs, where

$$\Psi_q(v) = \Phi_q(x, s, \mu) := \frac{e^T v^2 - n}{2} + \frac{e^T v^{1-q} - n}{q-1}, \quad 1 \le q \le 1 + \log n,$$

is used to define the SR search directions. Due to the specific choice of the SR-proximity function, system (1.16) has the form given in (2.7). The duality gap prediction for different search directions and different target μ values are discussed in Section 2.2.

In the sequel we present our SR-PC algorithm. At each iteration the algorithm has a predictor step and a corrector step. In the predictor step it makes either an adaptive SR step, or an affine scaling step in order to reduce the duality gap as much as possible while staying in the given large neighborhood. In the corrector step it recenters to a smaller neighborhood. In the predictor step the decrease of the duality gap for the SR and the affine scaling steps is compared. If the reduction of the duality gap for the affine scaling step is bigger than the one theory guarantees for the SR step, then it makes an affine scaling step, otherwise it makes an adaptive SR step. With this adaptive choice of the predictor step the best known polynomial iteration complexity of large-update SR-IPMs is preserved, and the SR-PC algorithm is enhanced with the quadratic convergence of the MTY predictor-corrector IPM, see Mizuno et al. (1993). The SR-PC algorithm is outlined as Algorithm 3.

Remark 4.2. In Algorithm 3 the step sizes α_1^* and α_3^* are the minimum step sizes warranted by theory, while α_1 and α_3 are the actual calculated step sizes at the given iterate.

Algorithm 3: Adaptive Large Neighborhood SR-PC-IPM

Input:

a proximity parameter $\tau > 1$ and $\eta_\ell(n,\tau) = \frac{n^\tau - n}{\log n}$;
an accuracy parameter $\epsilon > 0$, $1 \le q \le 1 + \log n$;
$(x,s) = (x^0, s^0)$ such that $\Phi_\ell(x,s,\mu_g) \le \eta_\ell(n,\tau)$;

begin

while $x^T s \ge \epsilon$ **do**

begin

Corrector step

solve (2.7) with $\mu = \mu_\ell^*$ and choose a step size α_1 such that
$\alpha_1 = \arg\min_{\alpha} \left\{ \Phi_\ell(x(\alpha), s(\alpha), \mu_\ell^*) \;\middle|\; (x(\alpha), s(\alpha)) \in \mathcal{N}_\ell(n,\tau) \right\}$;
Let $(x,y,s) = (x(\alpha_1), y(\alpha_1), s(\alpha_1))$.

end

begin

Predictor step

solve (4.2) and choose the maximum step size α_3 such that
$(x(\alpha_3), y(\alpha_3), s(\alpha_3)) \in \mathcal{N}_\ell(n,\tau)$;
if[a] $(1-\alpha_3) \le 1 - \alpha_3^* + \alpha_3^* \frac{(\mu_\ell^t)^{\frac{q+1}{2}}}{\mu_g (\mu_q^h)^{\frac{q-1}{2}}}$,
accept the affine scaling step, $\alpha = \alpha_3$;
else
solve (2.7) with $\mu = \mu_\ell^t$ derived from (4.3),
determine the maximum step size α_4 such that
$(x(\alpha_4), y(\alpha_4), s(\alpha_4)) \in \mathcal{N}_\ell(n,\tau)$, $\alpha = \alpha_4$;

end

end

Let $(x,y,s) = (x(\alpha), y(\alpha), s(\alpha))$.

end

[a] The value of α_3^* is given in Theorem 4.2.

4.1 The Corrector Step

In this subsection we present the estimation of the decreasing of the proximity function in the corrector step when μ_ℓ^* (see Lemma 4.1) is the target value. Here we also give an upper bound for the step size that guarantees a sufficient reduction of the proximity function. Due to space limitations the proofs are omitted.

Theorem 4.1 (Salahi and Terlaky (2004b)). *Suppose that the present iterate is in the neighborhood $\mathcal{N}_\ell(n,\tau)$ and $(\Delta x, \Delta y, \Delta s)$ is the solution of (2.7) with $\mu = \mu_\ell^*$ and $1 \le q \le 1 + \log n$. Then, for*

$$\alpha_1^* = \frac{\theta_2(\tau) v_{\min}}{4\sigma_\ell \log n \exp(\frac{-q+1+\log n}{2})} \quad \text{and} \quad \alpha_2^* = \frac{\theta_2^2(\tau)}{16 \log n \exp(-q+1+\log n)}$$

we have

$$\Phi_\ell(x(\alpha_1^*), s(\alpha_1^*), \mu_\ell^*) \le \Phi_\ell(x, s, \mu_\ell^*) - \max\{\alpha_1^* \sigma_\ell^2,\ \alpha_2^* v(\alpha)_{\min}^{-\log n}\}, \qquad (4.4)$$

where $\theta_2(\tau) = \min\left\{(\exp(\tau-1)-2)^{\frac{1}{2}}, \exp(1-\tau)(1-\exp(-\frac{1}{2}))\right\}$.

4.2 The Predictor Step

In this subsection we present the behavior of the search direction for the different μ values that SR-PC algorithm chooses in the predictor step. It uses $\mu = 0$ (affine scaling) whenever the reduction of the duality gap is at least as much as theory predicts for the SR step with $\mu = \mu_\ell^t$. If the reduction is not satisfactory, the algorithm makes an SR step with $\mu = \mu_\ell^t$. In what follows we present the result that specify how the step size for the affine scaling step and for the SR step is computed for $\mu = \mu_\ell^t$. The step size selection rule also guarantees that after each iteration the proximity function is bounded by a prescribed value that is related to the definition of the neighborhood.

Theorem 4.2 (Salahi and Terlaky (2004b)). *Let $(\Delta x, \Delta y, \Delta s)$ be the solution of system (2.7), where $\mu = \mu_\ell^t$ is defined by equation (4.3) and $1 \le q \le 1 + \log n$. Then, for the step size* $\alpha_3^* = \dfrac{\theta_2(\tau)\exp(-\tau)}{4\sigma_\ell \log n \exp\left(\frac{-q+1+\log n}{2}\right)}$,

the relation

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_\ell^t) \leq \Phi_\ell(x, s, \mu_\ell^t) - \frac{1}{2}\alpha_3^* \sigma_\ell^2$$

holds.

Theorem 4.3 (Salahi and Terlaky (2004b)). *Let* $(\Delta x, \Delta y, \Delta s)$ *be the solution of system (2.7), where* $\mu = \mu_\ell^t$ *is defined by equation (4.3) and* $1 \leq q \leq 1 + \log n$. *Then for the step size* α_3^*, *the relation*

$$\Phi_\ell(x(\alpha_3^*), s(\alpha_3^*), \mu_g(\alpha_3^*)) \leq \eta_\ell(n, \tau) \tag{4.5}$$

holds.

Remark 4.3. If we make an affine scaling step in the predictor step then the duality gap reduction is at least as big as the warranted duality gap reduction of the SR step, so for the worst case iteration complexity it suffices to consider that in all iterations we are making a SR step.

Theorem 4.4 (Salahi and Terlaky (2004b)). *After at most*

$$\mathcal{O}\left(\exp\left(\frac{1 + \log n - q}{2}\right)\sqrt{n}\log n \log\frac{(\exp(\tau - 1) + 1)n}{\epsilon}\right)$$

iterations Algorithm 3 terminates with a solution for which $x^T s \leq \epsilon$.

Corollary 4.1. *If* $q = 1 + \log n$, *then Algorithm 3 stops after at most*

$$\mathcal{O}\left(\sqrt{n}\log n \log\frac{(\exp(\tau - 1) + 1)n}{\epsilon}\right)$$

iterations.

Corollary 4.2. *If* $q = 1$, *then Algorithm 3 stops after at most*

$$\mathcal{O}\left(n \log n \log\frac{(\exp(\tau - 1) + 1)n}{\epsilon}\right)$$

iterations.

4.3 Quadratic Convergence

In this section we prove the quadratic convergence of Algorithm 3 (see Salahi and Terlaky (2004b)). For monotone Linear Complementarity Prob-

lems (LCP), Ye and Anstreicher (Ye and Anstreicher (1993)) proved that in a predictor-corrector algorithms one has

$$|\Delta x_i \Delta s_i| = \mathcal{O}(\mu_g^2),\ i = 1, \ldots, n, \tag{4.6}$$

when μ_g is sufficiently small. Since LO is a special case of monotone LCP, relation (4.6) is valid for LO as well. Thus we have

$$|d_{x_i} d_{s_i}| = \frac{\Delta x_i \Delta s_i}{\mu_g} = \mathcal{O}(\mu_g).$$

Since Algorithm 3 has multiple choices in the predictor step and it has been proved that close to optimality the affine scaling step is dominant to the SR step, thus it suffices to prove the quadratic convergence for the affine scaling direction. The following lemma gives a lower bound for the step size in the predictor step.

Lemma 4.2 (Salahi and Terlaky (2004b)). *Let (x^k, y^k, s^k) be an iterate in Algorithm 3. If the present duality gap is so small that (4.6) holds, then the step size α in the predictor step satisfies $\alpha \geq 1 - O(\mu_g)$.*

From the proof of Lemma 4.2 (see Salahi and Terlaky (2004b)), for the next iterate $(x^{k+1}, y^{k+1}, s^{k+1})$ it follows that

$$\mu_g^{k+1} = (1 - \alpha)\mu_g^k = \mathcal{O}\left((\mu_g^k)^2\right).$$

Our quadratic convergence result is the following.

Theorem 4.5 (Salahi and Terlaky (2004b)). *Let the iterate (x^k, y^k, s^k) be generated by Algorithm 3. When μ_g is sufficiently small, the algorithm is quadratically convergent, and any accumulation point of the iterate is a strictly complementary optimal solution of the problem.*

5 SR-Proximity Based Infeasible IPMs

In this section we present a SR-infeasible IPM (IIPM) based on $\Gamma_{13}(t)$ as a kernel function, see Salahi et al. (2003). The dynamic algorithm idea of Peng and Terlaky (Peng and Terlaky (2002)) has been used in Salahi et al. (2003) to develop a SR-proximity based IIPM.

For given $x, s > 0$, the Newton direction for (1.2) in our IIPM is determined by the following linear system of equations:

$$\begin{aligned} A\Delta x &= -r_b, \\ A^T\Delta y + \Delta s &= -r_c, \\ s\Delta x + x\Delta s &= \mu e - xs, \end{aligned} \tag{5.1}$$

where r_b and r_c are residuals defined by

$$\begin{aligned} r_b &:= Ax - b, \\ r_c &:= A^T y + s - c. \end{aligned}$$

Note that, due to the specific choice of the kernel function $\psi(t) = \Gamma_{13}(t)$, we can rewrite the Newton system as

$$\begin{aligned} A\Delta x &= -r_b, \\ A^T\Delta y + \Delta s &= -r_c, \\ s\Delta x + x\Delta s &= \mu^2 x^{-1}s^{-1} - xs. \end{aligned} \tag{5.2}$$

Let us denote by $(\Delta x(\mu), \Delta y(\mu), \Delta s(\mu))$ the solution of system (5.2). By using Corollary 2.3 one can show that if the targeted parameter is μ_3^h (see Lemma 2.2), then the search direction based on our specific SR-proximity function and the standard Newton direction will predict the change of the duality gap in the same way, i.e.,

$$(x + \alpha\Delta x(\mu^h))^T(s + \alpha\Delta s(\mu^h)) = x^T s\left(1 - \alpha + \frac{\mu^h\alpha}{\mu_g} + \alpha^2\frac{\Delta x^T\Delta s}{x^T s}\right).$$

Another important factor in designing the algorithm is how to define the neighborhood. In the sequel we introduce the SR-proximity based neighborhood that is different from the one used in feasible IPMs. The definition of the SR-infeasible neighborhood is as follows

$$\mathcal{N}_{\mathcal{I}}(\tau,\beta) := \Big\{(x,y,s) \ : \ \Psi_3(v) \le \frac{(\tau-1)n}{2}, \ \|r_b\| \le \|r_b^0\|\frac{\mu_g}{\mu^0}\beta, \ \|r_c\| \le \|r_c^0\|\frac{\mu_g}{\mu^0}\beta\Big\}, \tag{5.3}$$

where (x^0, y^0, s^0) is an arbitrary triple with $x^0, s^0 > 0$, $\mu^0 = \mu_g^0$, and $\beta \ge 1$

so that the initial point (x^0, y^0, s^0) is in the neighborhood $\mathcal{N}_{\mathcal{I}}(\tau, \beta)$. Here we impose that $\tau \geq 10$, that is necessary for deriving some estimations in the complexity analysis. Salahi et al. (Salahi et al. (2003)) show that if $(x, y, s) \in \mathcal{N}_{\mathcal{I}}(\tau, \beta)$, then infeasibility is bounded by a multiple of μ and by the initial infeasibility.

The IIPMs that are based on the infeasible neighborhood $\mathcal{N}_{\mathcal{I}}(\tau, \beta)$ and that explore the idea of dynamic IPMs given in Peng and Terlaky (2002), are presented in Algorithm 4. Since the complexity analysis of the algorithm is analogue to the algorithm for feasible IPMs, we omit the details here. The following theorem give the iteration complexity for the SR-IIPM algorithm.

Theorem 5.1. *Let $\tau \geq 10$, and $t_0 = \max\left(1, \frac{\|(r_b^0, r_c^0)\|}{\mu_0}\right)$. Then the SR-IIPM algorithm will terminate after at most*

$$\mathcal{O}\left(n^2 \log \frac{n(\tau+1)\beta t_0}{\epsilon}\right)$$

iterations with a solution satisfying $x^T s \leq \epsilon$ and $\|(r_b, r_c)\| \leq \epsilon$.

Although the new algorithm does not improve the iteration complexity, it is comparable to some other softwares i.e., LIPSOL and OSL. For the computational results see Salahi et al. (2003).

6 Non-SR Functions

SR-proximity based IPMs give us almost $\sqrt{n}$ reduction of the worst case iteration complexity of large neighborhood IPMs. Motivated by these results Bai et al. (Bai and Roos (2004), Bai et al. (2003a) and Bai et al. (2003b)) identify some new kernel functions that are not SR, but share some common properties with them, and provide similar polynomial complexity bounds as given for SR-IPMs. They modify the SR conditions from Definition 2.1 in order to make the analysis of the algorithm easier. They also show that large classes of SR kernel functions $\Gamma_{p,q}(t)$ satisfy the modified conditions.

Although the analysis based on the new kernel functions proposed by Bai et al. (Bai and Roos (2004), Bai et al. (2003a) and Bai et al. (2003b)) is much simpler than the original analysis for general SR functions, it is not applicable for such general class of kernel functions as SR proximity based complexity analysis does. Here we list the non-SR kernel functions

Algorithm 4: SR-IIPM Algorithm

Input:
proximity parameters $\tau \geq 10$ and $\beta \geq 1$;
neighborhood $\mathcal{N}_{\mathcal{I}}(\tau, \beta)$;
an accuracy parameter $\epsilon > 0$;
$(x^0, y^0, s^0) \in \mathcal{N}_{\mathcal{I}}(\tau, \beta)$;
begin
while $\max\{x^T s, \|r_b\|, \|r_c\|\} \geq \epsilon$ **do**
begin
if $\frac{\mu_g}{\mu^h} \geq \frac{\tau}{2}$ then $\mu := \mu^h$;
else
$\mu := \mu_t$ is a root of (3.1) where $q = 3$,
end;
solve system (5.2) for $(\Delta x, \Delta y, \Delta s)$;
begin
determine a step size α such that[a]
$\Phi(x(\alpha), s(\alpha), \mu_t) \leq \Phi(x, s, \mu_t) - \frac{\alpha^*}{2}\Phi(x, s, \mu_t)$
and $(x(\alpha), y(\alpha), s(\alpha)) \in \mathcal{N}_{\mathcal{I}}(\tau, \beta)$;
$x := x + \alpha\Delta x$;
$y := y + \alpha\Delta y$;
$s := s + \alpha\Delta s$,
end
end
end

[a] The value of α^* is given in Corollary 6.4 (Salahi et al. (2003)).

proposed in Bai and Roos (2004), Bai et al. (2003a) and Bai et al. (2003b), and some of their properties. The interested reader can find the detailed analysis of IPMs based on non-SR kernel functions in Bai et al. (2003a) and Bai et al. (2003b).

In Bai et al. (2003a), Bai et al. replace condition (2.1) from Definition 2.1 by the following set of conditions

$$\psi'''(t) < 0, \tag{6.1}$$

$$2\psi''(t)^2 - \psi'(t)\psi'''(t) > 0, \qquad t < 1, \tag{6.2}$$

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, \qquad t > 1,\ \beta > 1. \tag{6.3}$$

$\psi_i(t)$	Iteration Bound
$t-1+\frac{t^{1-q}-1}{q-1}$	$\mathcal{O}(qn\log\frac{n}{\epsilon})$
$\frac{t^2-1}{2}+e^{\frac{1}{t}-1}-1$	$\mathcal{O}(\sqrt{n}\log^2 n\log\frac{n}{\epsilon})$
$\frac{t^2-1}{2}-\int_1^t e^{\frac{1}{\zeta}-1}d\zeta$	$\mathcal{O}(\sqrt{n}\log^2 n\log\frac{n}{\epsilon})$
$\frac{t^2-1}{2}+\frac{e^{\gamma(1-t)}-1}{\gamma}$	$\mathcal{O}(\sqrt{n}\log n\log\frac{n}{\epsilon})$

Table 1: *Non-SR kernel functions ($q>1$, $b>0$)*

Note that condition (6.2) is also satisfied for $t \geq 1$ since $\psi^{'}(t) \geq 0$ and $\psi^{'''}(t) \leq 0$ when $t \geq 1$, and that SR.2 and (6.2) are conditions on the barrier behavior of $\psi(t)$. Condition (6.3) only deals with $t > 1$. Thus, it concerns the growth behavior of $\psi(t)$. The connection between conditions (6.1) and (6.3) is given by the following lemma.

Lemma 6.1 (Bai at al. (2003a)). *If $\psi(t)$ satisfies (6.1) and*

$$t\psi^{''}(t) - \psi^{'}(t) \geq 0,$$

then $\psi(t)$ satisfies (6.3).

The non-SR kernel functions introduced in Bai and Roos (2004), Bai et al. (2003a) and Bai et al. (2003b), are listed in the first column of Table 1. In the second column of Table 1 are given the iteration complexities of IPMs based on these functions. The first three listed functions in Table 1 satisfy conditions (1.10)–(1.12) and (6.1)–(6.3), and permit good iteration bounds. The first kernel function given in Table 1 differs from all other kernel functions since its growth term (i.e., $t-1$) is linear in t. This function was first introduced and analyzed in Bai and Roos (2004). The iteration complexity of large-update IPMs based on that function for specific q, is increasing with the increase of q, i.e., the complexity is $O(qn\log\frac{n}{\epsilon})$. Note that if $q = O(1)$, then this iteration bound is the same as the bound for the logarithmic barrier function.

The second function in Table 1 is the limit of the following sequence of functions

$$\psi_k(t) = \frac{t^2-1}{2} + (1+\frac{1}{k})^{1-k}\left((1+\frac{1}{kt})^k - (1+\frac{1}{k})^k\right), \quad k = 1, 2, \ldots.$$

By using Lemma 2.1.2 from Peng et al. (2002b), one can show that $\psi_k(t)$ is a SR function for every $k \geq 1$. Furthermore, for any fixed $t > 0$, one has

$$\lim_{k \longrightarrow \infty} \psi_k(t) = \frac{t^2 - 1}{2} + e^{\frac{1}{t} - 1} - 1.$$

This result implies that this non-SR function is the limit point of a sequence of SR functions. Hence, the cone of SR functions is not a closed cone. Note that the large-update algorithm based on the second non-SR function from Table 1, has the iteration bound $\mathcal{O}(\sqrt{n} \log^2 n \log \frac{n}{\epsilon})$.

The third and the second functions from Table 1, are closely related. Namely, if we integrate the approximation for the exponential part of the second function, we have a sequence of SR functions that converge to the exponential part of the third function. Therefore, this function is on the boundary of the cone of SR functions. The iteration complexity of large-update IPMs based on that function is $\mathcal{O}(\sqrt{n} \log^2 n \log \frac{n}{\epsilon})$, that is a factor $\log n$ weaker than the best complexity bound of large-update IPMs based on the $\Gamma_{1 \log n}(t)$ SR-kernel function.

In Bai et al. (2003b), Bai et al. define a family of kernel functions having finite values at the boundary of the feasible region, i.e., these are not barrier functions. This property distinguishes them from all other kernel functions to date. For $\gamma > 0$, the last function in Table 1 belongs to that family. Here we list the properties of that function:

1. $\lim_{t \to 0} \psi(t) = \psi(0) < \infty$,
2. $\psi''(t) > 1 \quad \text{for} \quad t > 0$,
3. $\psi'(t) + t\psi''(t) > 0 \quad \text{for} \quad t \geq \frac{1}{\gamma}$.

The first property, specify the finite value of the function at the boundary of the feasible region. Because of the second condition, we say that function $\psi(t)$ is uniformly strictly convex, see Peng et al. (2001). The third condition guarantees that $\psi(t)$ satisfies the SR.2 condition on the interval $[\frac{1}{\gamma}, \infty)$, see Lemma 2.1.

We point out that the forth function in Table 1, for $\gamma > 0$, is in a way a first order approximation of the following SR function:

$$\Gamma_{1q}(t) = \frac{t^2 - 1}{2} + \frac{t^{-q+1} - 1}{q - 1},$$

where $q = 1 + \gamma$. Namely,

$$\Gamma_{1(1+\gamma)}(t) = \frac{t^2-1}{2} + \frac{t^{-\gamma}-1}{\gamma} = \frac{t^2-1}{2} + \frac{e^{(-\gamma \log t)}-1}{\gamma} \approx \frac{t^2-1}{2} + \frac{e^{\gamma(1-t)}-1}{\gamma},$$

where the approximation in the exponent is due to the first order Taylor expansion

$$\log t \approx t - 1,$$

of the logarithm. Following this idea, we believe that other finite non-SR kernel functions can be derived as an approximation of SR functions.

7 Primal-dual IPMs for SDO Based on SR–Proximities

Semidefinite Optimization (SDO) is an extension of LO where the nonnegativity constraints are replaced by positive semidefiniteness on the matrix variables. SDO has a wide range of applications in combinatorial optimization (Goemans (1997), Lasserre (2002) and Zhao (1998)) and in control theory (Boyd et al. (1994), Fares et al. (2002) and Vandenberghe and Boyd (1996)), as well as in engineering fields (Anjos and Vanelli (2002), Laurent (2002), Li et al. (2002) and Vandenberghe and Boyd (1996)). Practice shows that semidefinite models for combinatorial optimization problems are sometimes significantly stronger than purely linear ones. Nowadays, IPMs provide a most powerful approach for solving SDO problems. Most IPMs for SDO can be viewed as natural extensions of the IPMs for LO, resulting with similar polynomial complexity results. Here we generalize the approach of the previous sections to the case of SDO, see Peng et al. (2002b).

We consider the SDO problem in the form

$$\begin{array}{lll} \min \ \mathrm{Tr}(CX) & & \text{(PSDO)} \\ \mathrm{Tr}(A_j X) & = & b_j, \quad j = 1, \ldots, m, \\ X & \succeq & 0, \end{array}$$

where C and A_j $(1 \leq j \leq m)$ are symmetric $n \times n$ matrices, $b := (b_1, \ldots, b_m)^T \in I\!R^m$, and $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix. The notation '$X \succeq 0$' (resp. '$X \succ 0$') means that X is symmetric positive semidefinite (resp. positive definite). The matrices A_i are assumed to be linearly independent.

Let S be a dual slack matrix. Then, the dual problem of (PSDO) is

$$\begin{aligned} \max\ & b^T y \qquad\qquad\qquad\qquad\qquad\qquad (\text{DSDO}) \\ \sum_{j=1}^{m} y_j A_j + S &= C, \\ S &\succeq 0, \end{aligned}$$

where $y \in I\!R^m$.

The concept of the central path can also be generalized from LO to SDO. This was first done by Nesterov and Nemirovski (Nesterov and Nemirovski (1994)) by introducing the extended logarithmic barrier function $\log\det(X)$ for the positive semidefinite constraint. Independently, Alizadeh (1995) applied IPMs to solving SDO problems arising from combinatorics. Followed by their approach, the *perturbed optimality conditions* are

$$\begin{aligned} \mathrm{Tr}(A_j X) &= b_j, \quad X \succeq 0, \qquad j = 1, \ldots, m, \\ \sum_{j=1}^{m} y_j A_j + S &= C, \quad S \succeq 0, \\ XS &= \mu I, \end{aligned} \tag{7.1}$$

where $\mu > 0$, and I denotes the $n \times n$ identity matrix. We may assume without loss of generality (see Klerk (2002) and Klerk et al. (1998)) that strict feasibility holds for both (PSDO) and its dual (DSDO). Under this assumption for every $\mu > 0$ there exists a unique solution $(X(\mu), y(\mu), S(\mu))$ of system (7.1), see Nesterov and Nemirovski (1994). The central path for SDO is defined by the set $\{(X(\mu), y(\mu), S(\mu)) : \mu > 0\}$. The basic idea of IPMs is to follow this central path and approach the optimal set of SDO by letting μ go to zero. Suppose that the point (X, y, S) is strictly feasible. Newton's method amounts to linearizing system (7.1), thus yielding the following equation:

$$\begin{aligned} \mathrm{Tr}(A_j \Delta X) &= 0, \qquad\qquad j = 1, \ldots, m, \\ \sum_{j=1}^{m} \Delta y_j A_j + \Delta S &= 0, \\ X\Delta S + \Delta X S &= \mu I - XS. \end{aligned} \tag{7.2}$$

A *crucial observation* for SDO is that the Newton system (7.2) might not have a symmetric solution ΔX. Many authors have suggested several ways

for symmetrizing the third equation in (7.2) so that the resulting new system has a unique symmetric solution, see e.g., Helmberg et al. (1996), Kojima (1997), Monteiro (1997) and Todd (1999). In Todd (1999), Todd analyzes more than twenty different search directions for SDO. Among others, the most popular directions are: the Alizadeh, Haeberly, Overton (**AHO**) direction introduced in Alizadeh et al. (1997), the search direction independently proposed by Helmberg, Rendl, Vanderbei and Wolkowicz (Helmberg et al. (1996)), and Kojima, Shindoh and Hara (Kojima et al. (1997)), and later rediscovered by Monteiro (Monteiro (1997)), which we refer to as the **H..K..M** direction, and the Nesterov and Todd (**NT**) direction (Nesterov and Todd (1997) and Nesterov and Todd (1998)).

Peng et al. (Peng et al. (2002b)) consider the symmetrization scheme from which the **NT** direction (Nesterov and Todd (1997), Nesterov and Todd (1998) and Todd et al. (1998)) is derived. The important reason for this is that the **NT** scaling technique transfers the primal variable X and the dual S into the same space; the so-called *V–space*. Let

$$P = X^{\frac{1}{2}}(X^{\frac{1}{2}}SX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}} = S^{-\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{\frac{1}{2}}S^{-\frac{1}{2}}, \tag{7.3}$$

and $D = P^{\frac{1}{2}}$, where for any symmetric positive definite matrix G, the exponent $G^{\frac{1}{2}}$ denotes its symmetric square root. Now, the matrix D can be used to rescale X and S to the same matrix V defined by (see Klerk (2002) and Sturm and Zhang (1999))

$$V := \frac{1}{\sqrt{\mu}}D^{-1}XD^{-1} = \frac{1}{\sqrt{\mu}}DSD. \tag{7.4}$$

Note that $D, V \succ 0$. Using the above notation one can state the *centrality condition* as $V = E$, and the *duality gap* as $\mu\mathrm{Tr}(V^2)$. Let us further define

$$\begin{aligned} \bar{A}_i &:= \frac{1}{\sqrt{\mu}}DA_iD, \qquad\qquad i = 1, \ldots, m, \\ D_X &:= \frac{1}{\sqrt{\mu}}D^{-1}\Delta XD^{-1}, \quad D_S := \frac{1}{\sqrt{\mu}}D\Delta SD. \end{aligned} \tag{7.5}$$

Then the *scaled* **NT** *search direction* $(D_X, \Delta y, D_S)$ is obtained from the

system (see also Nesterov and Todd (1997) and Sturm and Zhang (1999))

$$\begin{aligned} \mathrm{Tr}(\bar{A}_j D_X) &= 0, \qquad j = 1, \ldots, m, \\ \sum_{j=1}^{m} \Delta y_i \bar{A}_j + D_S &= 0, \\ D_X + D_S &= V^{-1} - V. \end{aligned} \tag{7.6}$$

The solution of system (7.6) is unique, and the unscaled direction can be easily derived from (7.5).

Here, analogous to the LO case, new search directions can be derived from SR kernel functions. Peng et al. (Peng et al. (2002a) and Peng et al. (2002b)) show that the **NT**-direction is induced by the kernel function of the logarithmic barrier function. Now we explain their approach and derive SR search directions based on SR-kernel functions.

First we provide some basic results of matrix analysis (Bellman (1995)). Let ψ be any kernel function. The definition of ψ can be extended to any diagonalizable matrix with positive eigenvalues, see Bellman (1995). In particular, given an eigen-decomposition

$$V = Q_V^{-1} \mathrm{diag}\left(\lambda_1(V), \lambda_2(V), \ldots, \lambda_n(V)\right) Q_V,$$

of V, with Q_V nonsingular,[1] the matrix function $\psi(V)$ is defined by

$$\psi(V) = Q_V^{-1} \mathrm{diag}\left(\psi(\lambda_1(V)), \psi(\lambda_2(V)), \ldots, \psi(\lambda_n(V))\right) Q_V. \tag{7.7}$$

Since $\psi(t)$ is twice continuously differentiable, the derivatives $\psi'(t)$ and $\psi''(t)$ are well-defined for $t > 0$. Hence, replacing $\psi(\lambda_i(V))$ in (7.7) by $\psi'(\lambda_i(V))$ and $\psi''(\lambda_i(V))$ respectively, for each i, we obtain the matrix functions $\psi'(V)$ and $\psi''(V)$.

Remark 7.1. Further on, when we use function $\psi(\cdot)$ and its derivatives $\psi'(\cdot)$ and $\psi''(\cdot)$, these denote matrix functions if the argument is a matrix and a univariate function if the argument is in $\mathbb{R}$.

We give now the notation of SR functions (Peng et al. (2002b)) on the cone of positive definite matrices $\mathcal{S}_{++}^{n\times n}$.

[1] The matrix Q_V is not unique, but $\psi(V)$ is well defined whenever $\psi(t)$ is well defined on the eigenvalues of V (Bellman (1995), p. 90). Note that since V is symmetric, we can choose Q_V to be orthogonal, i.e., $Q_V^{-1} = Q_V^T$.

Definition 7.1. A matrix function $\psi(V)$ given by (7.7) is SR on $\mathcal{S}_{++}^{n\times n}$ if the kernel function $\psi(t)$ is SR.

Peng et al. (Peng et al. (2002a) and Peng et al. (2002b)) define the *proximity measure* for SDO in the following way:

$$\Phi(X,S,\mu) := \Psi(V) := \mathrm{Tr}(\psi(V)) = \sum_{i=1}^{n} \psi(\lambda_i(V)), \tag{7.8}$$

where $\psi(V)$ is given by (7.7). Note that $\Psi(V) = 0$ if and only if V is the identity matrix, i.e., $XS = \mu I$, and otherwise $\Psi(V) > 0$, due to the properties of an SR-kernel function.

Remark 7.2. From now on, we assume that $\psi(\cdot)$ is SR. Hence, $\Phi(X,S,\mu)$ is a SR-proximity function.

Analogous to the LO case, using the NT scaling, the *new search direction* is given by the following system (see Peng et al. (2002a) and Peng et al. (2002b))

$$\begin{aligned} \mathrm{Tr}(\bar{A}_j D_X) &= 0, \qquad j = 1,\ldots,m, \\ \sum_{j=1}^{m} \Delta y_j \bar{A}_j + D_S &= 0, \\ D_X + D_S &= -\psi'(V). \end{aligned} \tag{7.9}$$

Having D_X and D_S, ΔX and ΔS can be calculated from (7.5). Note that the new search direction is a slight modification of the NT direction. The orthogonality of ΔX and ΔS follows from the orthogonality of D_X and D_S, i.e.,

$$\mathrm{Tr}(\Delta X \Delta S) = \mathrm{Tr}(D_X D_S) = \mathrm{Tr}(D_S D_X) = 0.$$

One may easily verify that if $\psi(t)$ is the kernel function of the logarithmic barrier function, then $\Phi(X,S,\mu)$ coincides with the classical primal-dual logarithmic barrier function for SDO, and the search direction is then precisely the **NT**-direction. In the next subsection we deal with properties of SR-proximity functions.

7.1 Properties of SR–Proximity Functions for SDO and the Algorithm

The following proposition gives properties of SR–proximity functions that are crucial for the analysis of SR–IPMs for SDO.

Proposition 7.1 (Peng et al. (2002b)). *Let the functions $\psi : \mathcal{S}_{++}^{n\times n} \to \mathcal{S}_{++}^{n\times n}$ and $\Psi : \mathcal{S}_{++}^{n\times n} \to I\!R$ be defined by (7.7) and (7.8), respectively. If $\psi(t)$ is a SR function, then*

(i) $\Psi(X)$ is strictly convex with respect to $X \succ 0$ and vanishes at its global minimal point $X = E$, i.e., $\Psi(E) = 0, \psi'(E) = 0_{n\times n}$. Further, there exist two constants $\nu_1, \nu_2 > 0$ such that

$$\nu_1(X^{p-1} + X^{-1-q}) \preceq \psi''(X) \preceq \nu_2(X^{p-1} + X^{-1-q}), \quad p, q \geq 1.$$

(ii) For any $X_1, X_2 \succ 0$,

$$\Psi\left([X_1^{\frac{1}{2}} X_2 X_1^{\frac{1}{2}}]^{\frac{1}{2}}\right) \leq \frac{1}{2}\left(\Psi(X_1) + \Psi(X_2)\right).$$

Statements *(i)* and *(ii)* of Proposition 7.1 can be viewed as transparent extensions of conditions SR1 and SR2 introduced in Section 2, respectively.

The following proposition provides some more properties of SR proximities for SDO.

Proposition 7.2 (Peng et al. (2002b)). *Let the kernel function $\psi(t)$ be SR, the proximity $\Psi(V)$ defined by (7.8), and*

$$\sigma := \sqrt{\mathrm{Tr}(\psi'(V)^2)} = \|\psi'(V)\|. \tag{7.10}$$

If the kernel function $\psi(t)$ satisfies condition SR1 (Definition 2.1, page 6), then

$$\Psi(V) \leq \frac{\sigma^2}{2\nu_1}, \tag{7.11}$$

$$\lambda_{\min}(V) \geq \left(1 + \frac{q\sigma}{\nu_1}\right)^{-\frac{1}{q}}, \tag{7.12}$$

$$\lambda_{\max}(V) \leq \left(1 + \frac{p\sigma}{\nu_1}\right)^{\frac{1}{p}}. \tag{7.13}$$

Statements (7.11)–(7.13) of Proposition 7.2 can be viewed as extensions of conditions derived for the case of LO (see Proposition 3.1.5, page 50 in Peng et al. (2002b)), where $v_i, v_{\max},$ and $v_{\min}$ are replaced by $\lambda_i(V), \lambda_{\max}(V)$, and $\lambda_{\min}(V)$, respectively.

The following corollary gives the relationships between the scaled duality gap $\|V\|^2$ and the proximity $\Psi(V)$.

Corollary 7.1. *Let the kernel function $\psi(t)$ be SR, and let the proximity $\Psi(V)$ be defined by (7.8). If the kernel function $\psi(t)$ satisfies condition SR1 (Definition 2.1, page 6), then*

$$\mathrm{Tr}(XS) = \mu\|V\|^2 \leq \mu\left(\sqrt{n} + \sqrt{\frac{2\Psi(V)}{\nu_1}}\right)^2. \tag{7.14}$$

Inequality (7.14) shows that the proximity yields an upper bound for the duality gap. Hence, it can be used as a potential function for minimizing the duality gap. The new algorithm for SDO (see Algorithm 5 and algorithm on page 118 in Peng et al. (2002b)) has the same structure as the algorithm for LO (see Algorithm 1 on page 213).

Remark 7.3. Algorithm 5 terminates with a point satisfying $n\mu < \epsilon$. By using (7.14), we obtain that

$$\mathrm{Tr}(XS) < \mu\left(\sqrt{n} + \sqrt{\frac{2\tau}{\nu_1}}\right)^2.$$

Hence if $\tau = \mathcal{O}(n)$, which means that the algorithm works indeed in a large neighborhood of the central path, then the algorithm finally reports a feasible solution such that $\mathrm{Tr}(XS) = \mathcal{O}(\epsilon)$.

7.2 Complexity of the Algorithm

As in the LO case, the key issue in the analysis of the algorithm is to estimate the decrease of the proximity during one step. In order to do that we introduce the following notation. Let $X_+ = X + \alpha\Delta X$, $S_+ = S + \alpha\Delta S$, and

$$V_+ = \frac{1}{\sqrt{\mu}} D^{-1} X_+ D^{-1} = \frac{1}{\sqrt{\mu}} D S_+ D,$$

Algorithm 5: Large-Update Primal-Dual Algorithm for SDO

Input:
a proximity parameter $\tau \geq \nu_1^{-1}$;
an accuracy parameter $\epsilon > 0$;
a fixed barrier update parameter $\theta \in (0,1)$;
$(X \succ 0,\ S \succ 0)$ and $\mu = 1$ with $\Phi(X,S,\mu) \leq \tau$;**begin**
while $n\mu \geq \epsilon$ **do**
begin
$\mu := (1-\theta)\mu$;
while $\Phi(X,S,\mu) \geq \tau$ **do**
begin
solve system (7.9) for $\Delta X, \Delta y, \Delta S$;
determine a step size α;
$X := X + \alpha\Delta X$;
$y := y + \alpha\Delta y$;
$S := S + \alpha\Delta S$, **end**
end
end

for $D = P^{\frac{1}{2}}$ and P as it is given in (7.3). It is trivial to verify that V_+^2 is unitary similar to the matrix $X_+^{\frac{1}{2}} S_+ X_+^{\frac{1}{2}}$ and thus to $(V+\alpha D_X)^{\frac{1}{2}}(V+\alpha D_S)(V+\alpha D_X)^{\frac{1}{2}}$. This further implies that the eigenvalues of V_+ are precisely the same as those of the matrix

$$\tilde{V}_+ := \left((V+\alpha D_X)^{\frac{1}{2}}(V+\alpha D_S)(V+\alpha D_X)^{\frac{1}{2}}\right)^{\frac{1}{2}}. \tag{7.15}$$

Since the proximity after one step is defined by $\Psi(V_+)$, it follows immediately from the definition of the proximity measure (7.8) that

$$\Psi(V_+) = \Psi\left(\tilde{V}_+\right). \tag{7.16}$$

The decrease in the proximity during one step is considered as a function of the step size α, i.e.,

$$f(\alpha) := \Psi(V_+) - \Psi(V) = \Psi(\tilde{V}_+) - \Psi(V). \tag{7.17}$$

First that we want to know is how far we can go along the search direction, i.e., the maximal value $\alpha_{\max}$ of the step size. Let

$$\bar{D}_X = V^{-\frac{1}{2}} D_X V^{-\frac{1}{2}}, \quad \bar{D}_S = V^{-\frac{1}{2}} D_S V^{-\frac{1}{2}}. \tag{7.18}$$

Since

$$V + \alpha D_X = V^{\frac{1}{2}}(I + \alpha \bar{D_X}) V^{\frac{1}{2}}, \quad V + \alpha D_S = V^{\frac{1}{2}}(I + \alpha \bar{D_S}) V^{\frac{1}{2}},$$

it follows that

$$V + \alpha D_X \succ 0 \;\Leftrightarrow\; E + \alpha \bar{D}_X \succ 0 \quad \& \quad V + \alpha D_S \succ 0 \;\Leftrightarrow\; E + \alpha \bar{D}_S \succ 0.$$

Thus, the maximal feasible step size is dependent on the eigenvalues of the matrices $\bar{D}_X$ and $\bar{D}_S$. The next result gives some estimates of the norms of the matrices $\bar{D}_X$ and $\bar{D}_S$, and thus the maximal feasible step size $\alpha_{\max}$ as well.

Lemma 7.1 (Peng et al. (2002b)). *Let the matrices $\bar{D}_X$ and $\bar{D}_S$ be defined by (7.18) and σ by (7.10). Let $\alpha_{\max}$ be the maximal feasible step size. Then*

$$||\bar{D}_X||^2 + ||\bar{D}_S||^2 \leq \bar{\alpha}^{-2}$$

and

$$\alpha_{\max} \geq \bar{\alpha},$$

where

$$\bar{\alpha} = \sigma^{-1}\left(1 + \frac{q\sigma}{\nu_1}\right)^{-\frac{1}{q}}.$$

Now we evaluate the function $f(\alpha)$ defined by (7.17). From part (ii) of Proposition 7.1 one gets

$$f(\alpha) = \Psi(\tilde{V}_+) - \Psi(V) \leq \frac{1}{2}\left(\Psi\left(V + \alpha D_X\right) + \Psi(V + \alpha D_S)\right) - \Psi(V) := f_1(\alpha).$$

Hence it suffices for us to estimate the decrease of the value of the function $f_1(\alpha)$ after one step. The main difficulty in the estimation of the function $f_1(\alpha)$ is to evaluate its first and second derivatives. In Peng et al. (2002b), Peng et al. compute

$$f_1'(\alpha) \;=\; \frac{1}{2}\mathrm{Tr}(\psi'(V + \alpha D_X) D_X + \psi'(V + \alpha D_S) D_S),$$

$$f_1''(\alpha) \;=\; \frac{1}{2}\frac{d^2}{d\alpha^2}\mathrm{Tr}(\psi(V+\alpha D_X)+\psi(V+\alpha D_S)),$$

and report the following result.

Lemma 7.2 (Peng et al. (2002b)). *If $\alpha < \bar{\alpha}$, then*

$$f_1''(\alpha) \le \frac{\nu_2\sigma^2}{2}\left((\lambda_{\max}(V)+\alpha\sigma)^{p-1}+(\lambda_{\min}(V)-\alpha\sigma)^{-q-1}\right).$$

The following corollary is a direct consequence of Lemma 7.2, and

$$f(0)=f_1(0)=0, \quad f'(0)=f_1'(0)=-\frac{\sigma^2}{2}.$$

Corollary 7.2. *If $\alpha < \bar{\alpha}$, then*

$$\begin{aligned} &f(\alpha) \le f_1(\alpha) \le \\ &-\frac{\sigma^2\alpha}{2}+\frac{\nu_2\sigma^2}{2}\int_0^\alpha\int_0^\xi\left((\lambda_{\max}(V)+\zeta\sigma)^{p-1}+(\lambda_{\min}(V)-\zeta\sigma)^{-q-1}\right)d\zeta d\xi. \end{aligned} \tag{7.19}$$

We define now the function $f_2(\alpha)$ as the right hand side of the inequality in (7.19). Obviously, $f_2(\alpha)$ is convex and twice differentiable for all $\alpha \in [0,\bar{\alpha})$. It is also easy to see that $f_2(\alpha)$ is decreasing at zero and that it goes to infinity as $\alpha \to \bar{\alpha}$. Let α^* be the point at which $f_2(\alpha)$ attains its global minimal value, i.e.,

$$\alpha^* = \arg\min_{\alpha\in[0,\bar{\alpha})} f_2(\alpha). \tag{7.20}$$

We can now state the following result.

Theorem 7.1 (Peng et al. (2002b)). *Let the function $f(\alpha)$ be defined by (7.17) with $\Psi(V) \ge \nu_1^{-1}$. Then the step-size α given by $\alpha = \alpha^*$ (7.20) or $\alpha = \nu_5\sigma^{-\frac{q+1}{q}}$ for*

$$\nu_5 := \min\left\{\frac{\nu_1}{2\nu_1\nu_2+p(\nu_1+2\nu_2)}, \frac{\nu_1^2}{(1+\nu_1)\left(2\nu_1\nu_2+q(\nu_1+2\nu_2)\right)}\right\},$$

is strictly feasible. Moreover,

$$f(\alpha) \leq \frac{1}{2} f'(0)\alpha \leq -\frac{\nu_5}{4} \left(\nu_1 \Psi(V)\right)^{\frac{q-1}{2q}} .$$

In the special case where $\psi(t) = \Upsilon_{p,q}(t)$*, this bound (with* $\nu_1 = \nu_2 = 1$*) simplifies to*

$$f(\alpha) \leq -\min\left(\frac{1}{12p+8}, \frac{1}{24q+16}\right) \Psi(V)^{\frac{q-1}{2q}}.$$

Since the proximity $\Psi(V)$ is determined by the eigenvalues of the matrix V, the growth behavior of the proximity $\Psi(V)$ is precisely the same as its LO counterpart $\Psi(v)$, see Peng et al. (2002b). If the current point enters the neighborhood again, then μ is updated to $(1-\theta)\mu$ for some $\theta \in (0,1)$. Proceeding as in the LO case, one can show that after the update of μ, the proximity is still bounded above by the number

$$\psi_0(\theta,\tau,n) := \frac{\nu_2 \tau}{\nu_1 (1-\theta)^{\frac{p+1}{2}}} + \nu_2 \Upsilon'_{p,q}\left((1-\theta)^{-\frac{1}{2}}\right) \sqrt{\frac{2n\tau}{\nu_1(1-\theta)}} + n\nu_2 \Upsilon_{p,q}\left((1-\theta)^{-\frac{1}{2}}\right). \quad (7.21)$$

Peng et al. (Peng et al. (2002b)) obtain the following bound.

Lemma 7.3 (Peng et al. (2002b)). *Let* $\Psi(X,S,\mu) \leq \tau$ *and* $\tau \geq \nu_1^{-1}$*. Then, after an update of the barrier parameter, no more than*

$$\left\lceil \frac{8q\nu_1^{-\frac{q-1}{2q}}}{\nu_5(q+1)} \left(\psi_0(\theta,\tau,n)\right)^{\frac{q+1}{2q}} \right\rceil$$

iterations are needed to recenter. In the special case when $\psi(t) = \Upsilon_{p,q}(t)$*, this bound (with* $\nu_1 = \nu_2 = 1$*) simplifies to*

$$\left\lceil \frac{8q \max(3p+2, 6q+4)}{q+1} \left(\psi_0(\theta,\tau,n)\right)^{\frac{q+1}{2q}} \right\rceil .$$

Finally, the complexity of the algorithm that is precisely the same as for the LO case, is stated by the following theorem.

Theorem 7.2 (Peng et al. (2002b)). *If $\tau \geq \nu_1^{-1}$, the total number of iterations required by the algorithm is not more than*

$$\left\lceil \frac{8q\nu_1^{-\frac{q-1}{2q}}}{\nu_5(q+1)} \left(\psi_0(\theta,\tau,n)\right)^{\frac{1+q}{2q}} \right\rceil \left\lceil \frac{1}{\theta} \log \frac{n}{\epsilon} \right\rceil.$$

In the special case where $\psi(t) = \Upsilon_{p,q}(t)$, this bound (with $\nu_1 = \nu_2 = 1$) simplifies to

$$\left\lceil \frac{8q \max(3p+2, 6q+4)}{q+1} \left(\psi_0(\theta,\tau,n)\right)^{\frac{q+1}{2q}} \right\rceil \left\lceil \frac{1}{\theta} \log \frac{n}{\epsilon} \right\rceil.$$

Hence, just as for LO, omitting the round off brackets in Theorem 7.2, by choosing $\theta \in (0,1)$, the iteration bound of the algorithm for SDO with large-updates is $\mathcal{O}(n^{\frac{q+1}{2q}} \log \frac{n}{\epsilon})$, while the algorithm with small-update ($\theta = \mathcal{O}(\frac{1}{\sqrt{n}})$) has an $\mathcal{O}(\sqrt{n} \log \frac{n}{\epsilon})$ iterations bound. Moreover, using Theorem 7.2 one can readily verify that if p is a small constant and $q = \log n$, then the new large-update algorithm has an $\mathcal{O}(\sqrt{n} \log n \log \frac{n}{\epsilon})$ iteration bound, the currently best bound for large-update methods.

8 Primal-Dual IPMs for SOCO Based on Self-Regular Proximities

Second-Order Conic Optimization (SOCO) is a generalization of LO and a specific case of SDO, and it holds the "sandwich relation" $LO \subset SOCO \subset SDO$. More precisely, SOCO is the problem of minimizing a linear objective function subject to the intersection of an affine set and the direct product of several second-order cones. A second-order cone in $\mathbb{R}^n$ can be defined by

$$K = \Big\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1^2 - \sum_{i=2}^{n} x_i^2 \geq 0, x_1 \geq 0\Big\}.$$

This cone is often referred to as the Lorentz cone in physics, but it is also known as "the ice-cream cone"[2]. Some general classes of problems, such as problems involving sums and maxima of norms, problems with hyperbolic

[2]This name comes from the similarity between the shape of a general second-order cone in the space $\mathbb{R}^3$ and the well known summer refreshment food.

constraints, matrix-fractional problems, and robust LO can be formulated as SOCO as well, see Lobo et al. (1998). Moreover, SOCO has a wide range of applications in areas like antenna array weight design (Lebret and Boyd (1997)), grasping force optimization (Cheng and Orin (1990)), FIR filter design (Cheney (1982) and Wu et al. (1998)), truss design (Achtziger et al. (1992) and Bendose el al. (1994)), etc. For details about different applications of SOCO see papers Ben-Tal and Nemirovski (2001) and Lobo et al. (1998). A SOCO problem can be solved by applying IPMs to the semidefinite formulation of a SOCO problem, or by applying IPMs *directly* which is showed to be a more efficient approach both in theory and practice (Andersen et al. (2003) and Nesterov and Nemirovski (1994)). Because of the inherent relations among LO, SOCO and SDO, most theoretical results for SOCO can be viewed as a transparent extension of LO, or a specialization of the results for SDO. Here we expand the approach of the previous sections to the case of SOCO, see Peng et al. (2002b) and Peng et al. (2002d).

We consider the SOCO problem in the following form:

$$\begin{array}{rcl} \min\ c^T x & & \\ Ax & = & b, \\ x & \succeq_K & 0, \end{array} \tag{PSOCO}$$

where K is the product of several second-order cones, i.e., $K = K^1 \times K^2 \ldots \times K^N$ with

$$K^j = \left\{ (x_1^j, \ldots, x_{n_j}^j)^T \in I\!R^{n_j} : \left(x_1^j\right)^2 \geq \sum_{i=2}^{n_j} \left(x_i^j\right)^2,\ x_1^j \geq 0 \right\},$$

$b \in I\!R^m$, $A \in I\!R^{m \times n}$ for $n = \sum_{j=1}^{N} n_j$, and $c, x \in I\!R^n$ where $x^T = \left((x^1)^T, (x^2)^T, \ldots, (x^N)^T\right)$, $x^j \in I\!R^{n_j}$, $j \in \mathcal{J} := \{1, 2, \ldots, N\}$. The notation $x \succeq_K s$ (respectively $x \succ_K s$) means that $x - s \in K$ (respectively $x - s \in K_+$, where K_+ denotes the interior of K). Here we assume that the matrix A is of full row rank, i.e., $\mathrm{rank}(A) = m$. Let $s \in I\!R^n$ be a dual slack variable. Then, the dual problem of (PSOCO) is

$$\max\ b^T y \tag{DSOCO}$$

$$\begin{array}{rcl} A^T y + s & = & c, \\ s & \succeq_K & 0, \end{array}$$

where $y \in I\!R^m$.

The concept of the central path for SOCO is very similar to the concept of the central path for SDO, and we refer to Klerk (2002), Sturm (1999) and Wolkowicz et al. (2000) for details. Let the operator $\mathrm{mat}(\cdot)$ be defined on $I\!R^{n_j}$, $j \in \mathcal{J}$ (see Peng et al. (2002b)) in the following way

$$\mathrm{mat}(x^j) := \begin{pmatrix} x_1^j & x_{2:n_j}^j \\ (x_{2:n_j}^j)^T & x_1^j I_{n_j-1} \end{pmatrix}, \quad x_{2:n_j}^j = (x_2^j, x_3^j, \ldots, x_{n_j}^j), \ x^j \in I\!R^{n_j}, \tag{8.1}$$

and the *generalized* operator $\mathrm{mat}(\cdot)$ on K as

$$\mathrm{mat}(x) := \begin{pmatrix} \mathrm{mat}(x^1) & \ldots & 0 \\ 0 & \ddots & 0 \\ 0 & \ldots & \mathrm{mat}(x^N) \end{pmatrix}, \quad x \in I\!R^n.$$

Then, the *perturbed optimality conditions* for SOCO are

$$\begin{array}{rcll} Ax & = & b, & x \succeq_K 0, \\ A^T y + s & = & c, & s \succeq_K 0, \\ \mathrm{mat}(x)s & = & \mu\tilde{e}, & \end{array} \tag{8.2}$$

where $\mu > 0$ and

$$\tilde{e}^T = \big((\tilde{e}^1)^T, \ldots, (\tilde{e}^N)^T\big), \quad (\tilde{e}^j)^T = \big(1, 0, \ldots, 0\big) \in I\!R^{n_j}, \ j \in \mathcal{J}.$$

Here we assume that both (PSOCO) and (DSOCO) satisfy the interior point condition, i.e., $x \succ_K 0$, $s \succ_K 0$. The central path for SOCO is defined by the solution sets $\{(x(\mu), y(\mu), s(\mu)) : \mu > 0\}$ of the system (8.2). It is easy to see that the linearized Newton system for (8.2) might not be well defined. To obtain a Newton–type system that has a *unique solution*, one usually refer to some scaling schemes. They were first proposed and studied by Tsuchiya (Tsuchiya (1997) and Tsuchiya (1998)). However, there are several popular choices for the scaling matrices and here we list some of them: the **AHO** search direction (Adler and Alizadeh (1995)); the primal (or dual) **H..K..M** direction, and the **NT** search direction (Tsuchiya (1997) and Tsuchiya (1998)). It can be shown that the linearized Newton system

in the scaled space induced by **NT** scaling is *always* well-defined if both $x \succ_K 0$ and $s \succ_K 0$ are feasible for SOCO.

Peng et al. (Peng et al. (2002b)) consider **NT** scaling for SOCO. One of the reasons for this is that for SOCO problems a large-update IPMs based on the **NT** search direction, always has a theoretically lower iteration bound than large-update IPMs relying on other search directions (Tsuchiya (1998)). Let

$$\bar{A} := \frac{1}{\sqrt{\mu}} A(U_{NT}W_{NT})^{-1}, \qquad v := \frac{1}{\sqrt{\mu}} U_{NT}W_{NT}x;$$
$$d_x := \frac{1}{\sqrt{\mu}} U_{NT}W_{NT}\Delta x, \qquad d_s := \frac{1}{\sqrt{\mu}} \left(U_{NT}W_{NT}\right)^{-1} \Delta s,$$

where W_{NT} is the scaling matrix, and U_{NT} is the diagonal matrix chosen such that (see Peng et al. (2002b), Tsuchiya (1997) and Tsuchiya (1998)) $U_{NT}W_{NT}x = (U_{NT}W_{NT})^{-1}s$. The existence of matrices W_{NT} and U_{NT} is proved in Peng et al. (2002b), Tsuchiya (1997) and Tsuchiya (1998). The important property of these matrices is that the gap between the primal and dual potential function in the scaled space, with respect to W_{NT} and U_{NT}, attains its global minimum value.

Note that $v \succ_K 0$. Using the above notation one can state the *centrality condition* as $v = \tilde{e}$. The **NT** search direction for SOCO is defined (see Andersen et al. (2003) and Tsuchiya (1998)) as the unique solution of the system

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T\Delta y + d_s &= 0, \\ d_x + d_s &= v^{-1} - v. \end{aligned} \tag{8.3}$$

In view of the orthogonality of Δx and Δs, one can easily verify that $d_x^T d_s = 0$. In the SOCO case, analogously to the SDO case, new search directions can be derived from SR kernel functions (Peng et al. (2002b) and Peng et al. (2002d)). Before we explain the approach of Peng et al. (2002b) and Peng et al. (2002d) and derive SR search directions based on SR-kernel functions, we provide some basic results about functions associated with a second order cone (Faraut and Korányi (1994), Faybusovich (1997) and Faybusovich (1998)).

Let $x^j \in I\!R^{n_j}$, $j \in \mathcal{J}$. Let $\lambda_{\max}(x^j)$ and $\lambda_{\min}(x^j)$ be the maximal and minimal eigenvalues respectively, of matrix $\mathrm{mat}(x^j)$ defined in (8.1). Let

$\psi(t)$ be any kernel function. Then the function $\psi : I\!R^{n_j} \to I\!R^{n_j}$ associated with the second-order cone K^j is defined (see Peng et al. (2002b) and Fukushima et al. (2001)) by

$$\psi(x^j) := \begin{cases} \left(\frac{1}{2}(\psi(\lambda_{\max}(x^j)) + \psi(\lambda_{\min}(x^j))), \frac{\psi(\lambda_{\max}(x^j)) - \psi(\lambda_{\min}(x^j))}{2||x^j_{2:n}||}(x^j)^T_{2:n}\right)^T, \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } x^j_{2:n} \neq 0, \\ (\psi(\lambda_{\max}(x^j), 0, \ldots, 0))^T, \qquad\qquad \text{if } x^j_{2:n} = 0, \end{cases} \tag{8.4}$$

and the *generalized* function $\psi : I\!R^n \to I\!R^n$ associated with the product of second-order cones $K = K^1 \times \ldots \times K^N$, is defined by

$$\psi(x) := ((\psi(x^1))^T, \ldots, (\psi(x^N))^T)^T, \tag{8.5}$$

where $x = ((x^1)^T, \ldots, (x^N)^T)^T$. It can be easily verified that if $\psi(t) \geq 0,\ t \geq 0$ and $x \in K$ then $\psi(x) \in K$ follows. Thus, it becomes clear that every nonnegative (positive) function on the nonnegative (positive) axis naturally extends to a function that maps (the interior of) K into itself.

Remark 8.1. Further on, when we use the function $\psi(\cdot)$ and its derivatives $\psi'(\cdot)$ and $\psi''(\cdot)$, these denote vector functions if the argument is a vector and a univariate function if the argument is in $I\!R$.

We give now the notation of SR functions on the product of second–order cones K.

Definition 8.1 (Peng et al. (2002b)). A function $\psi(x^j)$, associated with the second-order cone K^j, given by (8.4) is said to be SR if its kernel function $\psi(t)$ is SR. Analogous definition holds for the generalized function $\psi(x)$, that is given by (8.5).

From now on, we assume that $\psi(\cdot)$ is SR. The trace of a vector $x \in I\!R^n$ associated with the second-order cone K is defined as follows.

Definition 8.2. For any $x^j \in I\!R^{n_j}$, the *trace* of x^j associated with the second-order cone K^j is defined by

$$\mathrm{Tr}(x^j) = \lambda_{\max}(x^j) + \lambda_{\min}(x^j), \tag{8.6}$$

and for any $x \in I\!R^n$, $n = \sum_{j=1}^{N} n_j$, $x^T = ((x^1)^T, \ldots, (x^N)^T)$, $x^j \in I\!R^{n_j}$ the *generalized trace* of x associated with the product of second-order cones

$K = K^1 \times \ldots \times K^N$ is defined by

$$\mathrm{Tr}(x) = \sum_{j=1}^{N} (\lambda_{\max}(x^j) + \lambda_{\min}(x^j)).$$

Before describing the new search direction for SOCO, we need to define the SR–proximity measure used in the new IPMs for SOCO. Let $x^j, s^j \in I\!R^{n_j}$. Similar to the cases of LO and SDO, the SR–proximity measure for SOCO (Peng et al. (2002b)) is given by

$$\Psi(x^j, s^j, \mu) := \Psi(v^j) = \mathrm{Tr}(\psi(v^j)) = \psi(\lambda_{\max}(v^j)) + \psi(\lambda_{\min}(v^j)), \quad (8.7)$$

where $\psi(\cdot)$ is a univariate SR function, and for $x, s \in I\!R^n$, $n = \sum_{j=1}^{N} n_j$, $x^j \in I\!R^{n_j}$

$$\Psi(x, s, \mu) := \Psi(v) = \mathrm{Tr}(\psi(v)) = \sum_{j=1}^{N} \left(\psi(\lambda_{\max}(v^j)) + \psi(\lambda_{\min}(v^j))\right). \quad (8.8)$$

The last equality in (8.7) is derived from Definition 8.2 and from the properties of the kernel functions defined by (8.4) (see Lemma 6.2.6. in Peng et al. (2002b)). The *new search direction* proposed for SOCO is a slight modification of the **NT** direction and is defined, (see Peng et al. (2002b)) by the solution of the system

$$\begin{array}{lcl} \bar{A}d_x & = & 0, \\ \bar{A}^T\Delta y + d_s & = & 0, \\ d_x + d_s & = & -\psi'(v). \end{array} \quad (8.9)$$

Having d_x and d_s, Δx and Δs can be derived from (8.3). In the next subsection we discuss various properties of SR–proximities for SOCO.

8.1 Properties of SR–Proximity Functions for SOCO and the Algorithm

In order to provide basic properties of SR–proximity functions on second-order cones, we first define an algebra for the second–order cone.

Definition 8.3. The Euclidean Jordan algebra for second–order cone K^j

is defined by the bilinear operator

$$x^j \circ s^j = ((x^j)^T s^j, x_1^j (s_{2:n}^j)^T + s_1^j (x_{2:n}^j)^T)^T, \quad x^j, s^j \in I\!R^{n_j},$$

and for second–order cone $K = K^1 \times \dots K^N$ the Euclidean Jordan algebra is defined by

$$x \circ s = \left((x^1 \circ s^1)^T, \dots, (x^N \circ s^N)^T\right)^T, \quad x, s \in I\!R^n, n = \sum_{j=1}^{N} n_j,\ x^j, s^j \in I\!R^{n_j}.$$

Note that the Jordan product $\circ$ is commutative; for any $z \in I\!R^n$ one has $z \circ z \in K$, and for every $z \in K$ the equation $z \circ z = x$ has a unique solution z in K. It is easy to verify that

$$x \circ s = \mathrm{mat}(x)s = \mathrm{mat}(s)x = s \circ x,$$

and for any $x \in I\!R^n$

$$\mathrm{Tr}(x \circ x) = \sum_{j=1}^{N} 2||x^j||^2.$$

The following proposition characterizes several important properties of a SR functions associated with the second-order cone K.

Proposition 8.1 (Peng et al. (2002b)). *Let the functions $\psi(x) : K_+ \to K_+$ and $\Psi(x) : K_+ \to I\!R$ be defined by (8.5) and (8.8) respectively. If the function $\psi(x)$ is SR, then the following statements hold.*

(i) $\Psi(x)$ is strictly convex with respect to $x \in K_+$ and vanishes at its global minimal point $x = \tilde{e}$, i.e., $\Psi(\tilde{e}) = 0, \psi(\tilde{e}) = \psi'(\tilde{e}) = 0$. Further, there exist positive constants $\nu_1, \nu_2 > 0$ and $p, q \geq 1$ such that

$$\nu_1(x^{p-1} + x^{-1-q}) \preceq_K \psi''(x) \preceq_K \nu_2(x^{p-1} + x^{-1-q}); \tag{8.10}$$

(ii) Suppose x and s are two vectors in K_+. If $v \in K_+$ satisfies

$$\mathrm{Tr}(v^2) = \mathrm{Tr}(x \circ s)$$

and

$$\sum_{j=1}^{N} \lambda_{\max}((v^j)^2)\lambda_{\min}((v^j)^2) = \sum_{j=1}^{N} \lambda_{\max}(x^j)\lambda_{\max}(s^j)\lambda_{\min}(x^j)\lambda_{\min}(s^j),$$

then

$$\Psi\left(v\right) \leq \frac{1}{2}\left(\Psi(x) + \Psi(s)\right). \tag{8.11}$$

Comparing Proposition 8.1 with its SDO analogue Proposition 7.1 on page 248, we find that statements (ii) in these two propositions are slightly different. Actually, one can easily see that the matrix used in the second claim of Proposition 7.1 satisfies certain conditions similar to those posed in Proposition 8.1. However, the choice of the vector v allowing such conditions in second-order cones is much more strict. The following proposition gives some fundamental properties of SR–proximity functions that are crucial for the analysis of SR–IMPs for SOCO.

Proposition 8.2. *Let the kernel function $\psi(t)$ be SR, the proximity $\Psi(v)$ defined by (8.8), and*

$$\sigma^2 = \mathrm{Tr}(\psi'(v) \circ \psi'(v)) = 2\|\psi'(v)\|^2. \tag{8.12}$$

If the kernel function $\psi(\cdot)$ satisfies condition SR1 (Definition 2.1, page 6), then

$$\Psi(v) \leq \frac{\sigma^2}{2\nu_1}, \tag{8.13}$$

$$\lambda_{\min}(v) \geq \left(1 + \frac{q\sigma}{\nu_1}\right)^{-\frac{1}{q}}, \tag{8.14}$$

$$\lambda_{\max}(v) \leq \left(1 + \frac{p\sigma}{\nu_1}\right)^{\frac{1}{p}}. \tag{8.15}$$

Note the similarities of statements (8.13)–(8.15) in Proposition 8.2 and statements (7.11)–(7.13) in Proposition 7.2 on page 248, which are extensions of conditions derived for the case of LO (see Proposition 3.1.5 on page 50 in Peng et al. (2002b)). The following corollary gives the relationship between the duality gap and the proximity.

Corollary 8.1. *Let the kernel function $\psi(t)$ be SR, and let the proximity $\Psi(v)$ be defined by (8.8). If the kernel function $\psi(\cdot)$ satisfies condition SR1 (Definition 2.1, page 217), then*

$$\mathrm{Tr}(x \circ s) = 2\mu||v||^2 \leq 2N\mu + 4\mu\sqrt{\frac{N\Psi(v)}{\nu_1}} + \frac{2\Psi(v)}{\nu_1}\mu. \tag{8.16}$$

From Corollary 8.1 it follows that $\mathrm{Tr}(x \circ s) = \mathcal{O}(N\mu)$, whenever $\Psi(v) = \mathcal{O}(N)$. Hence, as in both the LO and the SDO case, the proximity is appropriate to use for minimizing the duality gap as μ goes to zero.

With respect to the exchange of the variables (e.g., semidefinite matrices with vectors associated with the second order cone), the new algorithm for SOCO has the same structure as the new algorithm for SDO (see Algorithm 5, page 249) and LO (see Algorithm 1, page 3). Analogues to the SDO case (see Remark 7.3), the following conclusion holds for the SOCO case.

Remark 8.2. The algorithm will stop when an iterate satisfies $N\mu < \epsilon$. By using (8.16), we obtain

$$x^T s = \frac{1}{2}\mathrm{Tr}(x \circ s) \leq N\mu + 2\mu\sqrt{\frac{N\tau}{\nu_1}} + \mu\frac{\tau}{\nu_1},$$

where $\tau \geq \nu_1^{-1}$ is a proximity parameter. For instance, for $\tau = N$ and the proximity satisfying condition SR1 with $\nu_1 = 1$, the algorithm works in a large neighborhood of the central path. One can easily verify that the algorithm will finally report a solution satisfying $x^T s \leq 4\epsilon$.

8.2 Complexity of the Algorithm

Like in the LO and the SDO cases, a crucial step in the estimate of the algorithm's complexity is to evaluate how fast one can reduce the value of the proximity for a feasible step along the search direction. Note that once the search direction $(\Delta x, \Delta s)$ is obtained, we need to decide how far we can go along this direction while staying in the feasible region. This amounts to estimating the maximal feasible step size. It should be noticed that for any step size α,

$$(x + \alpha\Delta x, s + \alpha\Delta s) \;\text{ is feasible }\; \Longleftrightarrow \;\; (v + \alpha d_x, v + \alpha d_s) \;\text{ is feasible.}$$

In the sequel we give a certain sufficient condition for a step size to be strictly feasible, and thus provide a *lower bound* for the maximal step size. The following lemma is proved in Peng et al. (2002b).

Lemma 8.1 (Peng et al. (2002b)). *Let* $\alpha_{\max}$ *be the maximal feasible step size and*

$$\bar{\alpha} = \lambda_{\min}(v)\sigma^{-1}. \tag{8.17}$$

Then

$$\alpha_{\max} \geq \bar{\alpha} \geq \sigma^{-1}\left(1 + \frac{q\sigma}{\nu_1}\right)^{-\frac{1}{q}}.$$

In view of Lemma 8.1, it is clear we can use any $\alpha \in (0, \bar{\alpha})$ as a step size. Note that, after such a step, we get a new primal-dual pair $(x + \alpha\Delta x, s + \alpha\Delta s)$ or the scaled pair $(v + \alpha d_x, v + \alpha d_s)$ and then we need to use the **NT** scaling scheme to transform the primal and dual vectors to the same vector, which we denote by v^+. On the other hand, according to (8.8), the proximity after this step is defined as $\Psi(v^+)$. Let us denote the difference between the proximity before and after one step as a function of the step size, that is

$$g(\alpha) = \Psi(v^+) - \Psi(v). \tag{8.18}$$

The main task now is to study the decreasing behavior of $g(\alpha)$ for $\alpha \in [0, \bar{\alpha})$.

From part (ii) of Proposition 8.1 one gets

$$g(\alpha) = \Psi(v^+) - \Psi(v) \leq \frac{1}{2}\left(\Psi\left(v + \alpha d_x\right) + \Psi(v + \alpha d_s)\right) - \Psi(v) =: g_1(\alpha).$$

Hence it suffices for us to estimate the decrease of the value of the functions $g_1(\alpha)$ after one step. In Peng et al. (2002b), Peng et al. compute

$$\begin{aligned} g_1'(\alpha) &= \frac{1}{2}\mathrm{Tr}(\psi'(v + \alpha d_x) \circ d_x + \psi'(v + \alpha d_s) \circ d_s), \\ g_1''(\alpha) &= \frac{1}{2}\frac{d^2}{d\alpha^2}\mathrm{Tr}(\psi(v + \alpha d_x) + \psi(v + \alpha d_s)), \end{aligned}$$

and report the following result that plays a crucial role in establishing the polynomial complexity of the algorithm.

Lemma 8.2 (Peng et al. (2002b)). *For any $\alpha \in (0, \bar{\alpha})$ holds*

$$g_1''(\alpha) \leq \frac{\nu_2 \sigma^2}{2} \left((\lambda_{\max}(v) + \alpha\sigma)^{p-1} + (\lambda_{\min}(v) - \alpha\sigma)^{-q-1} \right).$$

The result stated in Lemma 8.2 shows a close analogy with the SDO case (see Lemma 7.2, page 252). The remaining statements for the SOCO case are very similar to the LO and SDO cases, and hence we just state the results and recall their SDO analogues.

Corollary 8.2. *For $\alpha < \bar{\alpha}$,*

$$\begin{aligned} &g(\alpha) \leq g_1(\alpha) \leq \\ &-\frac{\sigma^2 \alpha}{2} + \frac{1}{2}\nu_2 \sigma^2 \int_0^{\alpha} \int_0^{\xi} \left((\lambda_{\max} v + zeta\sigma)^{p-1} + (\lambda_{\min} v - \zeta\sigma)^{-q-1} \right) d\zeta d\xi. \end{aligned} \tag{8.19}$$

Here we recall the SDO analogue of Corollary 8.2, i.e., Corollary 7.2 on page 252. We define now the function $g_2(\alpha)$ as the right hand side of the inequality in (8.19). It is straightforward to verify that $g_2(\alpha)$ is strictly convex and twice differentiable for all $\alpha \in [0, \bar{\alpha})$. Let α^* be the unique global minimizer of $g_2(\alpha)$ in the interval $[0, \bar{\alpha})$, i.e.,

$$\alpha^* = \arg \min_{\alpha \in [0, \bar{\alpha})} g_2(\alpha), \tag{8.20}$$

or equivalently α^* is the unique solution of the equation

$$\frac{\nu_2}{p} (\lambda_{\max} v + \alpha^* \sigma)^p - \lambda_{\max} v^p) + \frac{\nu_2}{q} \left(\lambda_{\min} v - \alpha^* \sigma)^{-q} - \lambda_{\min} v^{-q} \right) = \sigma.$$

The following lemma gives the estimation of the value of α^*.

Lemma 8.3 (Peng et al. (2002b)). *Let the constant α^* be defined by (8.20). Suppose that $\Psi(v) \geq \nu_1^{-1}$ and $v_{\max} > 1$ and let*

$$\nu_5 = \min \left\{ \frac{\nu_1}{2\nu_2(p + \mu_1) + \nu_1(p-1)}, \frac{\nu_1^2}{(1+\nu_1)\left(2\nu_2(\nu_1 + q) + \nu_1 q\right)} \right\}. \tag{8.21}$$

Then

$$\alpha^* \geq \nu_5 \sigma^{-\frac{q+1}{q}}.$$

The following result estimates the decreasing value of the proximity.

Theorem 8.1 (Peng et al. (2002b)). *Let the function $g(\alpha)$ be defined by (8.18) with $\Psi(v) \geq \nu_1^{-1}$. Then the step size given by $\alpha = \alpha^*$ (see (8.20)) or $\alpha = \nu_5 \sigma^{\frac{q-1}{q}}$ is strictly feasible. Moreover,*

$$nng(\alpha) \leq \frac{1}{2} g'(0)\alpha \leq -\frac{\nu_5 \nu_1^{\frac{q-1}{2q}}}{4} \Psi(v)^{\frac{q-1}{2q}}. \tag{8.22}$$

nn

Note that Theorem 8.1 is similar to its SDO analogue Theorem 7.1, page 252. To get the total complexity result for the algorithm, we still need to describe the growth behavior of the proximity $\Psi(v)$. In Peng et al. (2002b), Peng et al. show that after the update of μ, the proximity is still bounded above by the number $\psi_0(\theta, \tau, 2N)$ defined by (7.21), where N is the number of cones. The following result is proved in Peng et al. (2002b) (compare with its analogue Lemma 7.3).

Lemma 8.4 (Peng et al. (2002b)). *Let $\Psi(x, s, \mu) \leq \tau$ and $\tau \geq \nu_1^{-1}$. Then after an update of the barrier parameter, no more than*

$$\left\lceil \frac{8q\nu_1^{-\frac{q-1}{2q}}}{\nu_5 (q+1)} \left(\psi_0(\theta, \tau, 2N)\right)^{\frac{q+1}{2q}} \right\rceil$$

iterations are needed to recenter. In the special case where $\psi(t) = \Upsilon_{p,q}(t)$ with $\nu_1 = \nu_2 = 1$, at most

$$\left\lceil \frac{8q \max\{3p+1, 6q+4\}}{q+1} \left(\psi_0(\theta, \tau, 2N)\right)^{\frac{q+1}{2q}} \right\rceil$$

inner iterations are needed to recenter.

Finally, the total complexity of the algorithm can be estimated as follows.

Theorem 8.2 (Peng et al. (2002b)). *If $\tau \geq \nu_1^{-1}$, the total number of iterations required by the primal-dual Newton algorithm is not more than*

$$\left\lceil \frac{8q\nu_1^{-\frac{q-1}{2q}}}{\nu_5 (q+1)} \left(\psi_0(\theta, \tau, 2N)\right)^{\frac{q+1}{2q}} \right\rceil \left\lceil \frac{1}{\theta} \log \frac{N}{\epsilon} \right\rceil.$$

In the special case where $\psi(t) = \Upsilon_{p,q}(t)$ *with* $\nu_1 = \nu_2 = 1$*, the total number of iterations required by the primal-dual Newton algorithm is less than or equal to*

$$\left\lceil \frac{8q \max\{3p+1, 6q+4\}}{q+1} \left(\psi_0(\theta, \tau, 2N)\right)^{\frac{q+1}{2q}} \right\rceil \left\lceil \frac{1}{\theta} \log \frac{N}{\epsilon} \right\rceil .$$

Neglecting the influence of the constants in the expression in Theorem 8.2, one can safely conclude that for any fixed $\theta \in (0,1)$ with suitable $p, q \geq 1$, the algorithm with large-update for SOCO has an $\mathcal{O}(N^{\frac{q+1}{2q}} \log \frac{N}{\epsilon})$ iterations bound, while the algorithm with small-update ($\theta = \mathcal{O}(\frac{1}{\sqrt{N}})$) remains with the complexity of $\mathcal{O}(\sqrt{N} \log \frac{N}{\epsilon})$ iterations bound. Furthermore, from Theorem 8.2 one can readily see that if p is a constant and $q = \log N$, then the new large-update algorithm has a complexity $\mathcal{O}(\sqrt{N} \log N \log \frac{N}{\epsilon})$ iterations bound.

Finally, as we mentioned earlier a SOCO problem can also be solved by casting it as a SDO problem in $\mathcal{S}^{n\times n}$. In such a situation, the iteration complexity of the algorithm for solving the reformulated SDO problem has a bound of $\mathcal{O}(n^{\frac{q+1}{2q}} \log \frac{n}{\epsilon})$. If $K^j \in I\!R^2$ for all $j \in \mathcal{J}$, then the iteration complexity of the algorithm for SOCO is the same as that of its counterpart for SDO, but when $2N < n$, the algorithm that works directly on the original SOCO problem has a better iteration bound, see Peng et al. (2002b). However, the improvement is significant if $N \ll n$.

9 Summary and Future Work

9.1 Summary

The monograph Peng et al. (2002b) of Peng et al. presents the methodology of Self-Regular functions that provide a new framework for the theory of primal-dual IPMs. It contains their earlier research (Peng et al. (2002a), Peng et al. (2002c), Peng et al. (2002d) and Peng and Terlaky (2002)) that was a breakthrough in reducing the gap between the theory and the practical performance of IPMs with respect to small and large-update methods. In this paper we review the most recent developments on SR-IPMs for LO, and give an overview of SR-IPMs for SDO and SOCO problems as well. First, we describe an algorithmic schema of standard IPMs (see Algorithm

1, page 3), and than we explain the motivation for deriving the new search directions (see Section 1). The new IPMs use SR-functions as kernel functions in formulating proximity measures. We provide some basic properties of SR-functions and SR proximities in Section 2. The interested reader can find more about these functions in Peng et al. (2002b). In the remaining sections of this paper, we discuss several large-update path-following SR-IPMs for LO, SDO and SOCO, including their polynomial iteration complexity.

We present an adaptive single step large-update SR-IPM for LO including its complete complexity analysis in Section 3 (see also Salahi and Terlaky (2004a)). This method is developed for the $\Gamma_{1q},\ q > 1$ family of kernel functions. The adaptive large-update algorithm chooses the target value adaptively at each step, and hence it is more flexible in updating the target value than classical IPMs (see Algorithm 2, page 225). This algorithm retains the so far best $\mathcal{O}\left(\sqrt{n}\log n \log \frac{n}{\epsilon}\right)$ iteration complexity of large-update IPMs.

In Section 4, we present a SR-proximity based predictor-corrector IPM for LO that enjoys polynomial complexities and asymptotic superlinear convergence (see also Peng et al. (2003) and Salahi and Terlaky (2004b)). For the predictor step this algorithm chooses either a SR step or an affine scaling step. The corrector step is recentering with the respect to the SR-neighborhood. A remarkable feature is that the SR-neighborhood includes the infinity neighborhood that is usually used in predictor-corrector IPMs. Due to the specific step size strategy, this algorithm has the so-far best iteration complexity. Although feasible predictor-corrector IPMs admit better convergence results, infeasible IPMs (IIPMs) are widely used in academic and commercial softwares. In Section 5, we present a SR-IIPM that is based on the $\Gamma_{13}(t)$ function as a kernel function (see also Salahi et al. (2003)). The new IIPM always takes large-updates and consists of only one Newton step for each updated target value (see Algorithm 4, page 240). The iteration complexity for the SR-IIPM is $\mathcal{O}\left(n^2 \log \frac{n}{\epsilon}\right)$. The infeasible neighborhood which is used for designing the algorithm, is also defined in this section.

Recently, Bai et al. (Bai and Roos (2004), Bai et al. (2003a) and Bai et al. (2003b)) define new kernel functions that are not SR but attain similar iteration bounds as SR-IPMs do. They derive these functions by replacing the first condition SR.1 in the definition of SR functions, by a set of simpler

conditions (see conditions (6.1)–(6.3)), and therefore simplify the analysis of IPMs. In Section 6, we list the non-SR kernel functions, and give their iteration complexity. All of the presented functions are closely related to SR functions. We show that the non-SR function $\frac{t^2-1}{2} + e^{\frac{1}{2}-1} - 1$, is the limit point of a sequence of SR functions. Hence, it follows that the cone of SR functions is not a closed cone, and the exponential barrier functions are on the boundary of the SR-cone.

In Section 7 and Section 8, we explain the basics of SR-IPMs for SDO and SOCO, respectively. We show that the concept of SR-proximity based primal-dual IPMs for SDO is a natural extension of SR-IPMs for LO (see also Peng et al. (2002c)). Since the Newton system of the algorithm for SDO in general does not obtain a symmetric solution, some symmetrization schema should be applied. The algorithm given here, chooses the NT symmetrization schema (see Algorithm 5, page 250). The resulting large-update algorithm has an $\mathcal{O}(\sqrt{n}\log n \log \frac{n}{\epsilon})$ iteration bound which is up to date the best known one for SDO when large-update IPMs are applied. We show also that the derived SR-IPMs for SOCO admit similar complexity results as SR-IPMs for SDO do (see also Peng et al. (2002b) and Peng et al. (2002d)). The described algorithm for SOCO uses the NT scaling as well. Here we also provide the major results for SR-proximity functions for SOCO.

To sum up, all presented variants of IPMs, except IIPMs, improve the worst case iteration bound of large-update IPMs. Due to the remarkable results on improved theoretical complexity of IPMs, the theory of SR-IPMs is one of the hottest research area of IPMs.

9.2 Future Work

Theoretical and implementation aspects of SR-IMPs are still not fully explored. Here we list some of the challenging problems/questions that are worth to study in future research. Some of them are already under consideration.

1. Is it possible to design pure primal (or dual) IPMs for LO based on some barrier functions similar to the SR functions? If the result is positive, how to do it for SOCO and SDO problems?
2. Can we design SR-IPMs for SDO and SOCO based on the various

scaling techniques other than NT scaling?

3. Can we analyze SR-proximity based adaptive-large-update IPMs for SDO and SOCO?
4. Can we analyze SR-proximity based PC-IPMs for SDO and SOCO?
5. The analysis of SR-proximity based infeasible IPMs (IIPMs) was very complex for the LO case. The questions are:
   - How to generalize Algorithm 4 to the $\Gamma_{1q}(\cdot)$ family?
   - Can such generalization be combined with an adaptive update of the algorithm?
   - How to derive efficiently the generalization of IIPMs for SDO and SOCO?
6. How to generalize the analysis of IPMs based on non-SR functions to the whole family of non-SR functions that satisfy conditions (6.1)–(6.3)?
7. How to extend the SR-IPMs to Nonlinear Optimization?
8. Can we close the cone of SR functions?
9. Can we give a unified analysis for all kernel functions in the closure of the cone of SR functions?
10. Identifying other interesting sub–families and functions on the border of the SR cone.

Some preliminary numerical experiments with the new algorithms for LO and SOCO are given in Peng et al. (2003) and Zhu (2003). They show that the number of iterations of the SR-based IPMs is usually less than, or equal to that of the standard large-update IPMs. The preliminary results are promising, but still extensive numerical testing is needed to explore the efficiency of the approach. The following problems regarding the implementation should be addressed:

1. What choice of q and line search provide the best practical performance of SR based IPMs algorithms?

2. Implementation issues regarding SR-IPMs for LO still should be further explored.

3. Implementing SR-IPMs for SOCO and SDO should be fully considered.

References

Adler I. and Alizadeh F. (1995). Primal-Dual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems. Technical Report RRR–111–95, Rutcor, Rutgers Center for Operations Research, P.O. Box 5062, New Brunswick, NJ.

Achtziger W., Bendsoe M., Ben–Tal A. and J. Zowe (1992). Equivalent Displacement Based Formulations for Maximum Strength Truss Topology Design. *Impact of Computing in Science and Engineering* 4, 315–345.

Alizadeh F. (1995). Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization. *SIAM Journal on Optimization* 5, 13–51.

Alizadeh F., Haeberly J.-P.A. and Overton M.L. (1997). Complementarity and Nondegeneracy in Semidefinite Programming. *Mathematical Programming* 77, 111–128.

Andersen E.D., Gondzio J., Mészáros Cs. and Xu X. (1996). Implementation of Interior Point Methods for Large Scale Linear Programming. In: Terlaky T. (ed.), *Interior Point Methods of Mathematical Programming*. Kluwer Academic Publishers, 189–252.

Andersen E.D., Roos C. and Terlaky T. (2003). On Implementing a Primal-Dual Interior-Point Method for Conic Quadratic Optimization. *Mathematical Programming* 95, 249–277.

Anjos M. and Vanelli A. (2002). An Attractor-Repeller Approach to Floorplanning. *Mathematical Methods of Operations Research* 56, 3–27.

Bai Y.Q. and Roos C. (2004). A Primal-Dual Interior-Point Method Based on a New Kernel Function with Linear Growth Rate. *Proceedings Perth Industrial Optimization Meeting* (to appear).

Bai Y.Q., Roos C. and El Ghami M. (2003a). A Comparative Study of New Barrier Functions for Primal-Dual Interior–Point Algorithms in Linear Optimization. Working paper, 2003.

Bai Y.Q., Roos C. and El Ghami M. (2003b). A New Efficient Large-Update Primal-Dual Interior-Point Method Based on a Finite Barrier. *SIAM Journal*

on Optimization 13, 766–782.

Bellman R. (1995). *Introduction to Matrix Analysis.* Classics in Applied Mathematics, SIAM 12.

Bendose M., Ben–Tal A. and Zowe J. (1994). Optimization Methods for Truss Geometry and Topology Design. *Structural Optimization* 7, 141–159.

Ben-Tal A. and Nemirovski A. (2001). *Convex Optimization in Engineering: Modeling, Analysis, Algorithms.* MPS-SIAM Series on Optimization, SIAM.

Boyd S., El Gahaoui L., Feron E. and Balakrishnan V. (1994). *Linear Matrix Inequalities in System and Control Theory.* SIAM Studies in Applied Mathematics 15.

Cheng F.T. and Orin D.E. (1990). Efficient Algorithm for Optimal Force Distribution – the Compact–Dual LP Method. *IEEE Transactions on Robotics and Automation* 6, 178–187.

Cheney E.W. (1982). *Introduction to Approximation Theory.* Chelsea Publishing Company.

Faraut J. and Korányi A. (1994). *Analysis on Symmetric Cones.* Oxford University Press.

Fares B., Noll D. and Apkarian P. (2002). Robust Control via Sequential Semidefinite Programming. *SIAM Journal on Control and Optimization* 40, 1791–1820.

Faybusovich L. (1997). Euclidean Jordan Algebras and Interior-Point Algorithms. *Positivity* 1, 331–357.

Faybusovich L. (1998). A Jordan-Algebraic Approach to Potential-Reduction Algorithms. Technical Report, Department of Mathematics, University of Notre Dame, IN.

Fukushima M., Luo Z.Q. and Tseng P. (2001). Smoothing Functions for Second-Order-Cone Complementarity Problems. *SIAM Journal on Optimization* 12, 438–460.

Goemans M.X. (1997). Semidefinite Programming in Combinatorial Optimization. *Mathematical Programming* 79, 143–161.

Güler O. (1994). Limiting Behavior of the Weighted Central Paths in Linear Programming. *Mathematical Programming* 65, 215–228.

Güler O. and Ye Y. (1993). Convergence Behavior of Interior Point Algorithms. *Mathematical Programming* 60, 215–228.

Helmberg C., Rendl F., Vanderbei R.J. and Wolkowicz H. (1996). An Interior-Point Method for Semidefinite Programming. *SIAM Journal on Optimization* 6, 342–361.

Hung P. and Ye Y. (1996). An Asymptotically $O(\sqrt{n}L)$-Iteration Path-Following Linear Programming Algorithm that Uses Long Steps. *SIAM Journal on Optimization* 6, 570–586.

Jansen B., Roos C., Terlaky T. and Ye Y. (1997). Improved Complexity Using Higher Order Correctors for Primal-Dual Dikin Affine Scaling. *Mathematical Programming, Series B* 76, 117–130.

Karmarkar N.K. (1984). A New Polynomial-Time Algorithm for Linear Programming. *Combinatorica* 4, 373–395.

Klerk E. de (2002). *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*. Applied Optimization Series, 65. Kluwer Academic Publishers.

Klerk E. de, Roos C. and Terlaky T. (1998). Infeasible–Start Semidefinite Programming Algorithms via Self–Dual Embeddings. In: Pardalos P. and Wolkowicz H. (eds.), *Topics in Semidefinite and Interior–Point Methods*. American Mathematical Society, 215–236.

Kojima M., Shindoh S. and Hara S. (1997). Interior–Point Methods for the Monotone Semidefinite Linear Complementarity Problem in Symmetric Matrices. *SIAM Journal on Optimization* 7, 86–125.

Kojima M., Mizuno S. and Yoshise A. (1989). A Primal-Dual Interior Point Algorithm for Linear Programming. In: Megiddo N. (ed.) *Progress in Mathematical Programming: Interior Point and Related Methods*. Springer Verlag, 29–47.

Lasserre J.B. (2002). An Explicit Exact SDP Relaxation for Nonlinear 0-1 Programs. *SIAM Journal on Optimization* 12, 756–769.

Laurent M. (2003). Lower Bound for the Number of Iterations in Semidefinite Relaxations for the Cut Polytope. *Mathematics of Operations Research* 28, 871–883.

Lebret H. and Boyd S. (1997). Antenna Array Pattern Synthesis via Convex Optimization. *IEEE Transations on Signal Processing* 45, 526–532.

Li L., Luo Z.Q., Davidson T.N., Wong K.M. and Bosse E. (2002). Robust Filtering via Semidefinite Programming with Applications to Target Tracking. *SIAM Journal on Optimization* 12, 740–755.

Lobo M.S., Vandenberghe L., Boyd S. and Lebret H. (1998). Applications of Second-Order Cone Programming. *Linear Algebra and Its Applications* 284, 193–228.

Megiddo N. (1986). Pathways to the Optimal Set in Linear Programming. In: Megiddo N. (ed.) *Progress in Mathematical Programming: Interior Point and Related Methods*. Springer Verlag, 131–158. Also in Proceedings of the 6th

Mathematical Programming Symposium of Japan, Nagoya, Japan, 1986.

Mizuno S., Todd M.J. and Ye Y. (1993). On Adaptive Step Primal-Dual Interior-Point Algorithms for Linear Programming. *Mathematics of Operations Research* 18, 964–981.

Monteiro R.D.C., Adler I. and Resende M.G.C. (1990). A Polynomial-Time Primal-Dual Affine Scaling Algorithm for Linear and Convex Quadratic Programming and its Power Series Extensions. *Mathematics of Operations Research* 15, 191–214.

Monteiro R.D.C. (1997). Primal-Dual Path Following Algorithm for Semidefinite Programming. *SIAM Journal on Optimization* 7, 663–678.

Nesterov Y.E. and Nemirovski A.S. (1994). *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Studies in Applied Mathematics 13. SIAM.

Nesterov Y.E. and Todd M.J. (1997). Self-Scaled Barriers and Interior-Point Methods for Convex Programming. *Mathematics of Operations Research* 22, 1–42.

Nesterov Y.E. and Todd M.J. (1998). Primal-Dual Interior Point Methods for Self-Scaled Cones. *SIAM Journal on Optimization* 8, 324–364.

Peng J., Roos C. and Terlaky T. (2001). A New Efficient Large-Update Interior-Point Method for Linear Optimization. *Journal of Computational Technologies* 6, 61–80.

Peng J., Roos C. and Terlaky T. (2002a). A New Class of Polynomial Primal-Dual Methods for Linear and Semidefinite Optimization. *European Journal of Operations Research* 143, 234–256.

Peng J., Roos C. and Terlaky T. (2002b). *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Methods.* Princeton University Press.

Peng J., Roos C. and Terlaky T. (2002c). Self-Regular Functions and New Search Directions for Linear and Semidefinite Optimization. *Mathematical Programming* 93, 129–171.

Peng J., Roos C. and Terlaky T. (2002d). Primal-Dual Interior-Point Methods for Second-Order Conic Optimization Based on Self-Regular Proximities. *SIAM Journal on Optimization* 13, 179–203.

Peng J. and Terlaky T. (2002). A Dynamic Large-Update Primal-Dual Interior-Point Method for Linear Optimization. *Optimization Methods and Software* 17, 1077–1104.

Peng J., Terlaky T. and Zhao Y.B. (2003). A Predictor-Corrector Algorithm for Linear Optimization Based on a Specific Self-Regular Proximity Function. *SIAM Journal on Optimization* (to appear).

Renegar J. (2001). *A Mathematical View of Interior-Point Methods in Convex*

Optimization. SIAM.

Roos C., Terlaky T. and Vial J.-Ph. (1997). *Theory and Algorithms for Linear Optimization. An Interior Point Approach.* John Wiley.

Salahi M., Terlaky T. and Zhang G. (2003). The Complexity of Self–Regular Proximity based Infeasible IPMs. Technical Report 2003/3, Advanced Optimization Lab. Department of Computing and Software, McMaster University. `http://www.cas.mcmaster.ca/~oplab/publication`.

Salahi M. and Terlaky T. (2004a). An Adaptive Self-Regular Proximity Based Large-Update Interior Point Methods for LO. *Optimization Methods and Software* (to appear).

Salahi M. and Terlaky T. (2004b). An Adaptive Large Neighborhood Self-Regular Predictor-Corrector IPM for Linear Optimization. Submitted to *Journal of Optimization Theory and Applications.*

Sonnevend G. (1986). An "Analytic Center" for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming. In: Prékopa A., Szelezsán J. and Strazicky B. (eds.), *System Modeling and Optimization: Proceeding of the 12th IFIP Conference*, Lecture Notes in Control and Information Sciences, 84. Springer Verlag, 866–876.

Sturm J.F. (1999). Theory and Algorithms of Semidefinite Programming. In: Frenk H., Roos C., Terlaky T. and Zhang S. (eds.), *High Performance Optimization.* Kluwer Academic Publishers, 1–194.

Sturm J.F. and Zhang S. (1999). Symmetric Primal-Dual Path Following Algorithms for Semidefinite Programming. *Applied Numerical Mathematics* 29, 301–315.

Todd M.J. (1999). A Study of Search Directions in Primal-Dual Interior-Point Methods for Semidefinite Programming. *Optimization Methods and Software* 11, 1–46.

Todd M.J., Toh K.C. and Tütüncü R.H. (1998). On the Nesterov-Todd Direction in Semidefinite Programming. *SIAM Journal on Optimization* 8, 769–796.

Tsuchiya T. (1997). A Polynomial Primal-Dual Path-Following Algorithm for Second-Order Cone Programming. Technical Report 649, The Institute of Statistical Mathematics, Tokyo, Japan.

Tsuchiya T. (1999). A Convergent Analysis of the Scaling-Invariant Primal-Dual Path-Following Algorithms for Second-Order Cone Programming. *Optimization Methods and Software* 11/12, 141–182, 1999.

Vandenberghe L. and Boyd S. (1996). Semidefinite Programming. *SIAM Review* 38, 49–95.

Wolkowicz H., Saigal R. and Vandenberghe L. (2000). *Handbook of Semidefinite Programming (Theory, Algorithms and Applications).* Kluwer Academic Publishers.

Wright S.J. (1997). *Primal-dual Interior-Point Methods.* SIAM.

Wu S.P., Boyd S. and Vandenberghe L. (1998). FIR filter Design Via Spectral Factorization and Convex Optimization. In: Datta B. (ed.), *Applied and Computational Control, Signals and Circuits* 1. Birkhauser, 215–245.

Ye Y. and Anstreicher K.M. (1993). On Quadratic and $O(\sqrt{n}L)$ Convergence of a Predictor-Corrector Algorithm for LCP. *Mathematical Programming* 62, 537–551.

Ye Y. (1997). *Interior Point Algorithms, Theory and Analysis.* John Wiley.

Zhang Y. (1999). User's Guide to LIPSOL: Linear-Programming Interior Point Solvers V0.4. Interior Point Methods. *Optimization Methods and Software* 11/12, 385–396.

Zhao Q., Karisch S.E., Rendl F. and Wolkowicz H. (1998). Semidefinite Programming Relaxations for the Quadratic Assignment Problem. *Journal of Combinatorial Optimization* 2, 71–109.

Zhu X., Peng J., Terlaky T. and Zhang G. (2003). On Implementing Self-Regular Proximity Based Feasible IPMs. Technical Report 2003/2, Advanced Optimization Lab. Department of Computing and Software, McMaster University. `http://www.cas.mcmaster.ca/~oplab/publication.`

DISCUSSION

Jordi Castro

Universitat Politècnica de Catalunya, Spain

The gap between the theoretical running time and the practical performance of interior-point methods (IPMs) is one of their most noticeable features. State-of-the-art implementations, which are usually based on an infeasible IPM of complexity $O(n^2|\log\epsilon|)$, rarely need more than 100 iterations to obtain a solution, independently of the problem dimension. For instance, some difficult linear multicommodity problems of up to 1000000 variables and more than 300000 constraints are solved in Castro (2003a)

with an IPM in 200 iterations and about one day of execution (the dual simplex of CPLEX required more than 35 days of execution, and the IPM of CPLEX exhausted the 1Gb of memory of the computer). Even more dramatic is the number of iterations required by an IPM for the solution of some large convex separable quadratic multicommodity problems (i.e., 1000000 variables and 30000 constraints) arising in the field of statistical data protection: solutions are obtained in less than 10 iterations, Castro (2003b).

The excellent survey of Salahi, Sotirov and Terlaky discusses self-regular IPMs (SR-IPMs), a recent family of methods that reduce the theoretical complexity of current large-update—or long-step— feasible IPMs, from $O(n \log \frac{n}{\epsilon})$ to $O(\sqrt{n} \log n \log \frac{n}{\epsilon})$. For infeasible IPMs, which are used in practice, they provide a similar running time. That suggests some questions and comments, which are not addressed in the paper.

- The authors don't provide computational experience comparing infeasible SR-IPMs with current implementations (although some references are given). Current implementations of standard IPMs are based on Mehrotra's heuristic (Mehrotra (1992)) or the higher-order Gondzio corrections (Gondzio (1996)). Is there an equivalent heuristic for SR-IPMs to reduce the number of iterations performed?

- Even if a heuristic as the above mentioned is not available, I wonder if, from the better theoretical running time for the feasible case, we can expect a less number of iterations for implementations based on SR-IPMs instead of standard IPMs. And what about the execution times: are they comparable?

- For some very large-scale problems, we can not rely on Cholesky factorizations of the normal equations, and must use preconditioned conjugate gradients (PCG). Efficient preconditioners have been mainly devised for IPMs for network optimization problems (Castro (2000), Frangioni and Gentile (2004), Resende and Veiga (1993)). In these situations heuristic directions as Mehrotra's one are not effective, since they force the solution of two systems with the same matrix. The reduction in number of iterations is not worthwhile since the PCG must be applied twice (Castro (2000)). SR-IPMs may be a better alternative for this very large-scale problems that require PCG and pure (i.e., without heuristic directions) primal-dual IPMs. It seems

to be worth to explore the efficiency of SR-IPMs in these situations.

References

Castro J. (2000). A Specialized Interior-Point Algorithm for Multicommodity Network Flows. *SIAM Journal on Optimization* 10, 852–877.

Castro J. (2003a). Solving Difficult Multicommodity Problems Through a Specialized Interior-Point Algorithm. *Annals of Operations Research* 124, 35–48.

Castro J. (2003b). Quadratic Interior-Point Methods in Statistical Disclosure Control. Research Report DR2003-10, Dept. of Statistics and Operations Research, Universitat Politècnica de Catalunya. *Computational Management Science* (to appear).

Frangioni A. and Gentile C. (2004). New Preconditioners for KKT Systems of Network Flow Problems. *SIAM Journal on Optimization* 14, 894–913.

Gondzio J. (1996). Multiple Centrality Corrections in a Primal-Dual Method for Linear Programming. *Computational Optimization and Applications* 6, 137–156.

Mehrotra92 Mehrotra S. (1992). On the Implementation of a Primal-Dual Interior Point Method. *SIAM Journal on Optimization* 2, 575–601.

ResendeVeiga93 Resende M.G.C. and Veiga G. (1993). An Implementation of the Dual Affine Scaling Algorithm for Minimum-Cost Flow on Bipartite Uncapacitated Networks. *SIAM Journal on Optimization* 3, 516–537.

Tibor Illés

Eötvös Loránd University of Sciences, Hungary

Since the 1970's, one of the most intriguing research question in linear optimization (LO) has been the the following: is there an algorithm which solves any LO problem in *polynomial* (or *strongly polynomial*) number of iteration, i.e., the complexity of the algorithm is $\mathcal{O}(p(n, L))$ (or $\mathcal{O}(p(n))$), where p is a polynomial of finite degree, n is the number of variables and L is the input size of the LO problem.

The first part of the question was answered positively by Khacijan (1979) who showed that a variant of the *ellipsoid-method*, originally introduced by Yudin and Nemirovski (1976) for convex optimization, solves LO problems in polynomial number of iterations. The ellipsoid algorithm was studied intensively by several researchers both from theoretical and implementation point of view. In contrast to the good theoretical complexity of the algorithm, even the best known implementations of the ellipsoid algorithm is far from being competitive with the simplex-based LO solvers for practical problems. This was for the first time that an algorithm having the best worst-case theoretical complexity for a class of problems, had very poor practical efficiency in comparison with other algorithms having less attractive theoretical properties. Since all of the efforts to invent a practically efficient version of the ellipsoid algorithm for LO failed, most of the researchers lost their interest in dealing with this question. The new era started when Karmarkar (1984) published his *projective scaling primal interior point algorithm* for a special form of LO problems. It has been shown in Karmarkar's paper that every LO problem can be transformed into the form that projective scaling algorithm uses. The projective scaling algorithm enjoyed a polynomial complexity of $\mathcal{O}(nL)$ iterations and Karmarkar made the announcement that his algorithm could solve large-scale LO problems much faster than simplex-method based implementations. Nowadays, it is clear that Karmarkar opened a blooming, new field of modern interior point algorithms. Several researchers worked on introducing new and practically efficient variants of IPMs.

At the early stage, some authors observed that simple variants of the projective scaling algorithm could be tracked back to a very old nonlinear optimization algorithm known in the literature as *the logarithmic barrier method*. This observation led to a revival of several old methods from nonlinear optimization. Just two years after Karmarkar's publication, Sonnevend (1986) introduced the concept of *analytic center* and used the *central path* as a guideline to the optimal solution set. Sonnevend showed that his algorithms can be used to solve smooth convex optimization problems efficiently as well. This result linked interior point methods back to the class of smooth convex optimization problems. Megiddo (1986) recognized as well the importance of the central-path and studied its behavior. Almost all polynomial interior point algorithms follow the central path to reach the optimal solution set.

Primal-dual path-following methods where suggested by several authors at the end of the 1980's. Among others, Kojima et al. (1989) introduced a primal-dual Newton-type interior point algorithm for LO problems and later extended it for a new class of linear complementarity problems with the same polynomial complexity result. Theoretical properties of primal-dual type IPMs allow extensions of these algorithms to wider classes of convex optimization and complementarity problems. Alizadeh (1991) in his PhD thesis showed that semidefinite optimization problems can be efficiently solved using primal-dual IPMs both from theoretical and practical point of view. One main contribution to the theory of IPMs was made by Nesterov and Nemirovski (1994) who invented the theory of *self-concordant functions*, allowing some primal-dual type algorithm to be extended for nonlinear, smooth convex optimization; for classes of nonlinear complementarity problems; variational inequalities, and particularly to semidefinite optimization (SDO) and second-order conic optimization (SOCO), that have wide range of applications in combinatorial optimization, engineering etc. Comprehensive summary of IPMs for SDO can be found in de Klerk's PhD thesis (1997) and in the first part of his book.

All the above mentioned results provided the basis for the introduction of the concept of *self-regular functions* (SR), which is the central concept of Salahi, Sotirov and Terlaky's outstanding summary of the last generation interior point algorithms. The starting point of this research, that led to the definition of SR functions was the following observation: short step IPMs have better theoretical complexity, while long step IPMs are practically more efficient. What is the reason of this phenomena?

The first result, that tried to answer this question was published by Peng et al. (2000). The authors observed that one of the main component of the analysis of primal-dual Newton-type algorithm is based on the properties of the proximity measure used. In the past five years, Terlaky and his coauthors worked out several, important aspects of the theory of SR functions and their applications in analysis of primal-dual IPMs. First main result of their was that for a specific SR function, $\Gamma_{1,1+log\,n}$ used as proximity measure in analysis of primal-dual long-step IPMs, the best known complexity result for long-step IPMs were improved to $\mathcal{O}(\sqrt{n}\,\log n\,\log\frac{n}{\varepsilon})$.

Based on the research of Terlaky and his group several new variants of IPMs were developed for different classes of SR functions. Furthermore, conceptually new adaptive and long-step IPMs have been invented for LO,

SDO and SOCO.

The paper by Salahi, Sotirov and Terlaky presents a thorough treatment of self-regular IPMs for LO, SDO and SOCO. A short introduction into the theory of self-regular functions and primal-dual IPMs is provided. The paper discusses a rich amount of relevant literature which provides the reader with deeper insight into this research area and to actual research questions. However, those who want to learn more about IPMs and self-regular IPMs should consult books on IPMs written by Roos et al. (1997) and Peng et al. (2002).

Finally, I would like to mention again that most of the IPMs solve the LO problem in a polynomial number of iterations, but these algorithms are not strongly polynomial since they depend on the input size of the problem, as well. Therefore, the question which was asked in 1970's whether there exist polynomial or strongly polynomial algorithm to solve LO problems, is only partially answered, even after two decades of intensive research. We still do not know any strongly polynomial algorithm that solves general LO problems. However, it is more or less clear that any strongly polynomial algorithm for LO problem can be only pivot-type algorithm, since its complexity will depend (mainly) on the size of the problem. After the decades of IPMs whether the research in LO will focus again on pivot-type algorithms, I really do not know.

References

Alizadeh F. (1991). Combinatorial Optimization with Interior Point Methods and Semi-Definite Matrices. Ph.D. Dissertation, University of Minnesota, Minneapolis, USA.

Judin D.B. and Nemirowvski A.S. (1976). Informational Complexity and Effective Methods for the Solution of Convex Extremal Problems (in Russian). *Ekonomika i Matematcheskie Metody* 12, 357-369. (English translation: *Matekon: Translations of Russian and East European Mathematical Economics* 13, 25-45. Springer, 1977.)

Karmarkar N.K. (1984). A New Polynomial-Time Algorithm for Linear Programming. *Combinatorica* 4, 373-395.

Khacijan L.G. (1979). A polynomial Time Algorithm in Linear Programming. *Soviet Mathematics Doklady* 20, 191-194.

Klerk E. de (1997). Interior Point Methods for Semidefinite Programming. Ph.D.

Dissertation, Technical University of Delft, Delft, The Netherlands.

Kojima M., Mizuno S. and Yoshise A. (1989). A Primal-Dual Interior Point Algorithm for Linear Programming. In: Megiddo N. (ed.), *Progress in Mathematical Programming: Interior Point and Related Methods*. Springer Verlag, 29-47.

Megiddo N. (1986). Pathways to the Optimal Set in Linear Programming. In: Megiddo N. (ed.), *Progress in Mathematical Programming: Interior Point and Related Methods*. Springer Verlag, 131-158. Identical version in Proceedings of the 6th Mathematical Programming Symposium of Japan, Nagoya, Japan, 1-35, 1986.

Nesterov Y.E. and Nemirovski A.S. (1994). *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics 13. SIAM.

Peng J., Roos C. and Terlaky T. (2000). New Complexity Analysis of Primal-Dual Newton Method for Linear Optimization. *Annals of Operations Research* 99, 23-39.

Peng J., Roos C. and Terlaky T. (2002). *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Methods.* Princeton University Press.

Roos C., Terlaky T. and Vial J.-Ph. (1997). *Theory and Algorithms for Linear Optimization: An Interior Point Approach.* John Wiley.

Sonnevend Gy. (1986). An "Analytic Center" for Polyhedrons and New Classes of Global Algorithms for Linear (Smooth, Convex) Programming. In: Prékopa A., Szelezsán J. and Straziczky B. (eds.), *System Modelling and Optimization: Proceedings of the 12th IFIP-Conference held in Budapest, Hungary, September 1985*, Lecture Notes in Control and Information Sciences, 84. Springer Verlag, 866-876.

Florian Jarre

University of Düsseldorf, Germany

The development of modern interior-point methods for linear programs and classes of convex optimization problems has brought a fundamental change to the analysis and implementation of modern optimization algorithms.

For the first time there was a practically very efficient algorithm that has a polynomial complexity bound for linear optimization problems. The

numerical implementation of interior-point methods using sparse matrix technologies, efficient preprocessing strategies, and heuristics for the selection of certain parameters was tedious but very successful.

At the same time, the theoretical development of interior methods left two major questions:

1. Is there a more sophisticated analysis showing that less than order of $\sqrt{n}$ iterations are necessary to reduce the duality gap by a factor of 2?

2. The best complexities can be proved for so-called *short step variants* of interior-point methods which are numerically unattractive. Is it possible to prove the same (or an even better) complexity analysis for those *long step methods* that have been very efficient in practical implementations?

Improvements on the first question appear to be extremely difficult.

The contribution of the paper by Salahi, Sotirov and Terlaky is to help understanding the second question – and possibly some aspects of the first question as well. Recent results by Peng, Roos, Salahi, and Terlaky are summarized in this paper to form a unified theory for large step primal dual methods for linear as well as semidefinite and second order cone programming problems.

The key idea of the approach based on self-regular functions is to reformulate and relax the (bilinear) complementarity conditions using "suitable" functions ψ, namely self-regular functions. Here, "suitability" is ultimately determined by the behavior of Newtons method when applied to solve the (reformulated) primal dual system of optimality conditions. Exploiting the special structure of this primal dual system, the authors (based on the work of J. Peng) are able to identify sufficient conditions for the functions ψ to be "suitable". These conditions bound the second derivative of ψ from above and from below (relation (2.1) in the paper) and assume convexity of $\psi(e^t)$ (Lemma 2.1, Item 1.). As in the concept of self-concordance which has certainly motivated the name self-regularity – it turns out to be sufficient to identify conditions for a scalar function ψ and then "lift" these conditions to a separable function Ψ of the complementarity terms.

The authors give two examples of families of functions that satisfy the self-regularity conditions, state a long step interior method, and prove an

$O(\sqrt{n}\log n \log \frac{n}{\epsilon})$ bound for the total number of iterations needed to compute an ϵ approximation of the linear program. Moreover local superlinear convergence is also established. The given bound is the best complexity bound for large-update interior point methods and thus partially answers the second question mentioned above. "Partially" since the large-update methods that are currently implemented still differ from the method analyzed here.

An important generalization of the above concept is given in the second part of the paper starting with Chapter 7, where semidefinite and later also second-order cone programs are being considered. The search directions are based on a scaling matrix V that is difficult to compute. The NT-direction also depends on this scaling matrix, but can be computed without forming V. If it is possible to find numerical algorithms for evaluating the search directions analyzed in the present paper in an efficient way (without forming V), it is very likely that the extensions discussed here will be included in future updates of software packages for semidefinite optimization. The theoretical results regarding the number of iterations are the same as for linear programs. Moreover, for second-order cone programs the computation of the search directions seems numerically attractive. The paper concludes with several open questions, indicating that the computation of the search directions in semidefinite programming does not appear to be a topic of current research.

Summarizing, the paper forms a broad and well written survey on the most recent developments in the theory of large-update interior-point methods; promising first numerical tests of the concept are described in two recent preprints.

Arkadi Nemirovski
Technion – Israel Institute of Technology, Israel

The paper provides an in-depth overview of *Self-Regular Interior Point Methods* – a novel area in Interior Point science discovered by J. Peng, C. Roos and T. Terlaky in late 90's (see Peng et al. (2002a) and Peng et al. (2002b)) and rapidly developing, primarily due to the same authors, since then. The motivation and the importance of this line of research can be

explained as follows. At late 90's, the "common wisdom" said that the most attractive, both theoretically and computationally, IPMs for Linear, Second Order Conic and Semidefinite Programming are primal-dual path-following methods approaching the optimal solution by tracing what is called a primal-dual central path. As far as theoretical complexity bounds are concerned, the best of these methods are the "small-update" ones, travelling in a relatively narrow neighbourhood of the path and capable to solve the problem within accuracy ϵ in $O(1)\sqrt{n}\ln(n\epsilon^{-1})$ Newton-type iterations, where n is appropriately defined size of the problem (the number of linear/conic quadratic inequality constraints in Linear/Second Order Conic Programming, and the total row size of linear matrix inequality constraints in Semidefinite Programming). However, in actual computations the small-update IPMs are by far outperformed by their "large-update" counterparts which use much more "aggressive" stepsize policies and travel in much wider neighbourhoods of the central path, in spite of the fact that the theoretical complexity bounds for the latter methods are significantly larger ($O(1)n\ln(n/\epsilon)$ vs. $O(1)\sqrt{n}\ln(n/\epsilon)$) than those of small-update IPMs. The resulting discrepancy between theoretical performance guarantees and actual behaviour of IPMs remained for a long time a major "question mark" in Interior Point science. The theory developed by Peng, Roos and Terlaky resolved this challenging issue by discovering an essentially new family of large-update primal-dual methods for Linear, Second Order Conic and Semidefinite Programming, methods based on novel *self-regular* proximity measures. The complexity bounds for these new large-update methods can be made as low as $O(1)\sqrt{n}\ln(n)\ln(n/\epsilon)$, that is, essentially as good as the best known complexity bounds for small-update IPMs. I believe that the development of the Peng-Roos-Terlaky theory is *the* major event in Interior Point science during the last 5-7 years, and I fully agree with the claim (the very end of section 9.1) that the theory of Self-Regular IPMs "is one of the hottest research areas of IPMs".

The paper is a very good overview of the theory of self-regular IPMs, an overview which covers both "old" results (published in 2001 – 2002 – the basic theory of self-regular proximity measures and the associated large-update methods for Semidefinite and Second Order Conic Programming) and the most recent developments – primarily adaptive large-update algorithm and advanced predictor-corrector method for Linear Programming, as well as some other novelties. The presentation is clear and well-organized; this is not an easy reading, of course, but it hardly could be

so given the intrinsic complexity of the topic. In summary, I believe the authors did an excellent job by offering a wide audience a detailed presentation of one of the major recent developments in Interior Point science. In my opinion, an additional praise should be expressed for a feature which is not that common for survey-type publications, that is, for an excellent outline of challenging open research questions in the area (section 9.2). In a survey of a novel and rapidly developing research area, such an outline is really valuable.

All praises being said (and meant), I switch to the role of a "devil advocate" (which, I believe, is a part of commenter's duties). In light of the story presented in the survey, the situation with primal-dual IPMs can be summarized as follows. For the time being, there exist three families of methods of this type:

(i) small-update algorithms with the best known so far theoretical complexity bound $O(1)\sqrt{n}\ln(n\epsilon^{-1})$;

(ii) traditional large-update algorithms with worse complexity ($O(1)n\ln(n\epsilon^{-1})$) and much better practical performance than the one of small-update methods;

(iii) new self-regular large-update algorithms with the (nearly) best known so far complexity bound $O(1)\sqrt{n}\ln(n)\ln(n\epsilon^{-1})$.

My "devil's advocate" comment is that *by itself*, the above three facts do not say much about the advantages of the self-regular IPMs. What is missing is the comparison of *practical* performance of algorithms from families (iii) and (ii). Only a convincing evidence that the new methods are computationally at least as good as the traditional large-update primal-dual path-following algorithms could indeed justify the conclusion that the new developments eliminate the discrepancy between IPM theory and practice. Without this evidence, it is difficult to understand, e.g., why the self-regular IPMs should be treated as more attractive than the well-known potential reduction primal-dual methods (which share the complexity bound of the small-update methods and are, in a sense, "large update" techniques). Thus, we are in the situation where "the proof of the pudding is in eating", and it would be highly desirable for this kind of proof to be presented in the survey. To the best of my knowledge, the authors do have in their disposal a quite convincing computational evidence that (at least some of) the

methods discussed in the paper are as efficient in practice as the best of the traditional large-update IPMs. I am somehow surprised that all information on this subject contained in the survey reduces to a small paragraph at the end of section 9.2, with the summary that "the preliminary results are promising, but still extensive numerical testing is needed to explore the efficiency of the approach". I believe that this (already excellent) survey would gain, were the existing numerical experience presented in more details.

References

Peng J., Roos C. and Terlaky T. (2002a). *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Methods.* Princeton University Press.

Peng J., Roos C. and Terlaky T. (2002b). Self-Regular Functions and New Search Directions for Linear and Semidefinite Optimization. *Mathematical Programming* 93, 129–171.

Jiming Peng

McMaster University, Canada

I have to admit that I was quite hesitant when the editor first invited me to comment on a review of SR-IPMs, a research area discovered by myself during my PhD study, together with my two supervisors. Therefore, I might be a bit more critical than a general commenter.

This paper presents a survey about a special topic of interior-point methods (IPMs) based on the so-called self-regular (SR) proximities for solving classes of linear conic optimization problems. Since Karmarkar's ground-breaking paper (Karmarkar (1984)) in 1984, there have been tremendous works on IPMs and many variants of IPMs including IPMs in the pure primal space, the pure dual space and the scaled primal-dual space, feasible and infeasible IPMs, predictor-corrector type IPMs, have been well-studied (Roos et al. (1997), Wright (1997), Ye (1997)). Several efficient optimization packages based on IPMs have been developed (Andersen et al. (1996)). The IPM theory was matured around the middle of 1990s due to the celebrated self-concordant theory by Nemirovskii and Nesterov (1994). Since

then, the research on IPMs has shifted its focus to some particular issues such as warm-start and implementation heuristics while some open puzzles remains unanswered.

One of these puzzles concerns the gap between the theoretical complexity and practical performance of primal-dual large-update IPMs and the SR-approach is one of the major attempts within the IPM community to bridge this gap. The main framework of the SR approach was described in the monograph Peng et al. (2002), where feasible primal-dual IPMs based on SR-proximities were well-studied. The present paper covers not only the algorithms in Peng et al. (2002), but also their variants in the infeasible and predictor-corrector paradigm as well as their extensions to linear conic optimization and IPMs based new proximity functions that enjoy certain properties similar to that of SR functions. One particular feature that distinguishes SR-IPMs from its cousins of IPMs based on the classical Newton method is the adaptive version of the algorithms. A SR proximity, as a function of the duality gap parameter μ, might have a global minimum at some point that is less than a fraction $(x^T s/n)$ of the current duality gap. This partially explains why SR-IPMs and the corresponding analysis are very helpful in the study of large-update IPMs and it also makes the algorithm more attractive in implementation.

In spite of a broad coverage of various results about SR-approach, there are two important issues regarding SR-IPMs that have not been well addressed in this review. The first is the implementation of the algorithm and numerical test, which is documented in Zhu et al. (2003). In the corresponding package McIPM, a dynamic strategy is used. The dynamic strategy works as follows, it first tries the classical IPM and refers to SR-IPMs only when a step from the classical IPM failed to meet certain criteria. It has been observed that such a dynamic strategy is very helpful in dealing with some known relatively hard problems. In particular, the algorithm based on such a dynamic strategy does not fit into the variants discussed in this review and has not been explored yet in the literature.

The second issue concerns the design of SR-IPMs. At present, most SR-IPMs are built upon the bedrock of primal-dual framework, while the high-degree and non-logarithmic barrier in the proximity function plays a crucial role in the analysis. It is reasonable to expect that the employment of various nonlinear and non-logarithmic barrier functions will help in solving classes of nonlinear convex optimization problems for which the primal-dual

structure of the problem is not easy to construct. From a viewpoint of the algorithm design, the problem of designing SR-IPMs in the pure primal or dual space definitely deserves more attention than other problems listed in the future plan of the authors' research.

A minor point in the review is its organization. The present review did cover many results regarding SR-approach, and more appealing, it also provides a list of potential interesting research topics. However, except in the introduction where the authors described the motivation leading to SR-IPMs, in many sections of the review, the authors did not explain why we should have so many technical results. For example, in Section 2.2, the authors listed many properties of SR-proximities without any motivation and most of them are useful only for the analysis of adaptive version of the algorithm.

To summarize, the review has given a fairly good survey on SR-IPMs and illustrated that there exist still a lot of research opportunities in IPMs.

References

Andersen E.D., Gondzio J., Mészáros Cs. and Xu. X. (1996). Implementation of Interior Point Methods for Large Scale Linear Programming. In: Terlaky T. (eds.), *Interior Point Methods of Mathematical Programming*. Kluwer Academic Publishers, 189–252.

Karmarkar N.K. (1984). A New Polymonimal-Time Algorithm for Linear Programming. *Combinatorica* 4, 373–395.

Nesterov Y.E. and Nemirovskii A.S. (1994). *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics 13. SIAM.

Peng J., Roos C. and Terlaky T. (2002). *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms*. Princeton University Press.

Roos C., Terlaky T. and Vial J.-Ph. (1997). *Theory and Algorithms for Linear Optimization. An Interior Approach*. John Wiley.

Wright S.J. (1997). *Primal-Dual Interior-Point Methods*. SIAM.

Ye Y. (1997). *Interior Point Algorithms, Theory and Analysis*. John Wiley.

Zhu X., Peng J., Terlaky T. and Zhang G. (2003). On implementing self-regular proximity based feasible IPMs. Technical Report, Adv. Optimization Lab., Department of Computing and Software, McMaster University, Canada.

Robert J. Vanderbei[1]
Princeton University, USA

The paper *On Self-Regular IPMs* (SR-IPM) provides a detailed account of the theory of interior-point methods from the viewpoint of self-regular proximity functions. Included in this account is a complexity analysis of several variants of interior-point methods for linear optimization problems. Long-step variants with roughly $O(n)$ iteration bounds and short-step variants that are $\sqrt{n}$ better in their worst case behavior are described. The account begins with methods that assume a feasible initial solution but then later turns to the important "infeasible interior-point" methods in which the starting point for the algorithm need not be feasible.

After covering linear optimization, the discussion turns to semidefinite optimization and second-order cone optimization where again analogues of "the usual methods" are presented together with "the usual results". Not only does one find the usual results but also certain new methods that are inspired by a careful analysis of self-regular functions.

Focusing the theory on the notion of self-regular proximity functions provides what is certainly pedagogically the simplest and most elegant way to think about interior-point methods. It may also be true that the algorithmic variants that are inspired by this approach may prove to be "best in practice" if one could only convince the developers of the state-of-the-art codes to give these new methods a try. In any case, all workers in the field should familiarize themselves with this approach.

Speaking of pedagogy, I would like to offer a few words here about how I would improve (and least in my opinion) the account in one last way—since interior-point methods are fundamentally primal-dual algorithms, I would develop the theory in a primal-dual symmetric context. For linear optimization, this means that I would start by discussing primal problems in full-inequality form:

$$(P) \qquad \min\{c^T x : \ Ax \geq b, \ x \geq 0\}.$$

[1] The author gratefully acknowledges research support from the NSF (CCR-0098040) and the ONR (N00014-98-1-0036).

The dual problem is then the "negative-transpose" of the primal:

$$(D) \qquad -\min\{-b^T y : \ -A^T y \geq -c, \ y \geq 0\} = \max\{b^T y : \ A^T y \leq c, \ y \geq 0\}.$$

Before going on, let me anticipate and address some potential objections from pedagogues. First, one might argue that the equality form is closer to what one implements in practice and hence it is this form that should be considered. Furthermore, one might say that it is easy to convert from inequality form to equality form but harder to go the other way around, so again isn't it better to focus on equality form. To this I say that neither of these forms are what one encounters in practice. Rather, one finds a mixture of equality and inequality constraints (and even constraints with two-sided inequalities; i.e., inequalities with ranges) and variables that have general bounds on one side (say below), two sides (above and below), or no bounds at all. A practical implementation has to consider all of these possibilities. What simple version we choose to discuss really has little to do with practical implementations.

The second objection I envision (and have heard expressed) is that the inequality version involves two sets of slack variables and hence two sets of complementarity conditions and therefore makes for a more obtuse presentation. This is a valid pedagogical consideration. I would like to try to convince the reader that in fact the presentation can be further streamlined and thereby get more to the heart of the matter.

So, here we go. First we add slack variables to (P) and (D):

$$\begin{aligned} \min\{c^T x : \quad & Ax - r = b, \ x \geq 0, \ r \geq 0\} \\ \max\{b^T y : \quad & A^T y + s = c, \ y \geq 0 \ \ s \geq 0\}. \end{aligned}$$

Jumping straight to the infeasible interior-point method, system (38) in the SR-IPM becomes

$$\begin{aligned} A\Delta x - \Delta r &= -\rho_b \\ A^T \Delta y + \Delta s &= -\rho_c \\ s\Delta x + x\Delta s &= \mu e - xs \\ r\Delta y + y\Delta r &= \mu e - ry. \end{aligned}$$

This looks more complicated than (38), but, it's symmetry also suggests some streamlining of notation. In fact, we can rewrite this system as

$$\begin{aligned} B\Delta z - \Delta t &= -\sigma \\ t\Delta z + z\Delta t &= \mu e - zt \end{aligned}$$

where

$$B = \begin{bmatrix} 0 & -A^T \\ A & 0 \end{bmatrix}, \qquad z = \begin{bmatrix} x \\ y \end{bmatrix} \qquad t = \begin{bmatrix} s \\ r \end{bmatrix} \qquad \sigma = \begin{bmatrix} -\rho_c \\ \rho_b \end{bmatrix}.$$

Now, all of the different interior-point methods that arise from judiciously choosing a self-regular proximity function can be derived also in this context by applying these proximity functions to the complementarity constraint $t\Delta z + z\Delta t = \mu e - zt$.

For semidefinite optimization, a symmetric form of the problem is easy to state. The primal is

$$\min\{\mathrm{Tr}(CX) : \ \mathcal{A}X \succeq B, \ X \succeq 0\}$$

and the associated dual is

$$\max\{\mathrm{Tr}(BY) : \ \mathcal{A}^*Y \preceq C, \ Y \succeq 0\}.$$

Here, C and X are symmetric $n \times n$-matrices, B and Y are symmetric $m \times m$-matrices, $\mathcal{A}$ is a linear operator from $n \times n$-matrices to $m \times m$-matrices, and $\mathcal{A}^*$ is its adjoint. Note that this inequality formulation of the semidefinite optimization problem is not related in quite a trivial way to the usual equality formulation of the problem. Not only have we replaced an equality with an inequality, but we have also changed the dual space from m-vectors into symmetric $m \times m$-matrices. By enlarging the dual space, we have enlarged this class of optimization problems. Perhaps there are interesting "real world" problems that belong only to this larger class. If so, that alone would justify investigating to what extent the self-regularity idea can be used to develop interior-point methods for this class of problems. As far as I know, only two papers exist that deal with this larger class of problems: Vanderbei and Yang(1995) and Muramatsu and Vanderbei (1999).

Let me end by saying again that the paper SR-IPM provides an excellent survey of the very powerful and elegant self-regularity approach to the study of interior-point methods.

References

Muramatsu M. and Vanderbei R.J. (1999). Primal-Dual Affine-Scaling Algorithms Fail for Semidefinite Programming. *Mathematics of Operations Research* 24, 149–175.

Vanderbei and Yang (1995). The Simplest Semidefinite Programs Are Trivial. *Mathematics of Operations Research* 20, 590–596.

Rejoinder by M. Salahi, R. Sotirov, and T. Terlaky

The authors would like to thank the discussants for their valuable comments and remarks on the paper. We would also like to thank the discussants for their compliments regarding the clarity of explanations of the rather complex materials, and for the compliments regarding the organization of the paper.

Since most of the reviewers have asked for more details on computational and/or software issues, we will cover all their questions and discuss on that topic at the end of the rejoinder. We will also present illustrative computational results, and give references to the papers on implementations of SR–IPMs for those who are interested in more exhausting details.

In the sequel we reply to the reviewers on the other specific issues.

Reply to Jordi Castro

Please find the information regarding the computational implementation of SR-IPMs at the end of the rejoinder.

It is known that most of the algorithms implemented in IPM solvers use Mehrotra's heuristics (Mehrotra (1992)) and Gondzio's higher order

corrections (Gondzio (1996)). Mehrotra's predictor-corrector method has two components: an adaptive choice of the centering parameter, and then the computation of a second-order approximation of the central path. Implemented SR-proximity based predictor-corrector IPMs have the same two components as Methrotra's algorithm, but with SR-corrector search directions. The search directions are computed by solving either the normal equation system (using Cholesky factorization) or solving the augmented system (by using LDL^T factorization).

In the McIPM software package we do not use preconditioned conjugate gradients (PCG). Not only McIPM, but all high-performance IPM software is using a factorization of either the normal equation system or the augmented system (Andersen et al. (1996)). The primary reason is that the sparsity structure of those matrices do not change during the algorithm, and the costly reordering and symbolic factorization need to be done only once, before the algorithm starts. On the other hand, for PCG algorithms it is crucial to have good preconditioner. So far there is no efficient methodology to produce a good preconditioners, that has to change at each iteration. Successful applications of PCG algorithms in IPMs is restricted to specially structured network optimization (assignment, transportation, etc.) problems (Castro (2000) and Portugal et al. (1996)), when the normal equation by definition becomes a fully dense matrix. The other possible use of the PCG algorithm is when PCG is used in numerically refining (iterative refinement) the search directions after solving the normal equation systems (Gupta (2000)).

Reply to Florian Jarre.

We hope that the computational results for LO based on feasible IPMs presented at the end of the rejoinder and the details given in Zhu et al. (2003), answer completely the second question of the discussion paper.

Unfortunately there are no numerical results for SDO to date, but we do have promising preliminary ones for SOCO, Wang (2003). The main difficulty in the implementation of SR-proximity based IPMs for SDO is forming the matrix V indeed. If an efficient way of dealing with matrix V is not explored in the implementations, then the cost per step in SR-IPMs for SDO would be too high. This topic deserves future research and need to be added to the list of open topics.

Reply to Tibor Illés.

We would like to thank the discussant for an interesting discussion on algorithms which solves LO problems in polynomial number of iterations, in which he has embedded our SR–IPMs approach. Unfortunately, to date there is no strongly polynomial algorithm to solve LO problems. We share the discussant's hope that the rapid development of LO algorithms will finally give some strongly polynomial algorithms.

Reply to Arkadi Nemirovski.

We are glad that the discussant find important our outline of challenging open research questions regarding SR-IPMs. Our motivation to present these questions was to challenge the reader to join our effort to solve those research questions and to emphasize that SR–IPMs offer a rich area for challenging and fruitful research.

We discuss implementation issues at the end of this rejoinder. Computational evidences show that SR-IPMs should be implemented in new IPMs solvers.

Reply to Jiming Peng.

In the paper we have tried to limit our discussions only to those technical details that are important for understanding the background of each discussed variants of SR–IPMs, and the corresponding complexity analysis. In Section 2.2 we first give common properties of proximity functions that are based on the $\Gamma_{1q}(t)$ family for $q > 1$, and then separately the properties that apply only for the function $\Gamma_{1q}(t)$, where $q = 3$ and $q = 1 + \log n$.

Then we have presented an adaptive large–update SR-IPM for LO in more details, with the aim to present a complete self–contained analysis of an SR–IPM. For the other variants of SR–IPMs we present only the main results that could be easily absorbed once the contents from Section 2.2. is understood.

We also believe that the problem of designing SR–IPMs in the pure primal or dual space is a very important issue. However, it is far from being obvious how, if at all, it can be done, so it remains open for future

work.

Regarding the implementation issues mentioned in the discussion, we refer to the end of this rejoinder.

Reply to Robert J. Vanderbei.

We would like to thank the discussant for the idea on the improvement of the presentation of the SR approach for LO. It is true that practical implementations include both equality and inequality constraints, but we are concerned that using a formulation of a problem that include both inequality and equality constraints in the analysis, would additionally complicate the presentation of the variants of IPMs. The main goal of this paper is to introduce a "wide audience" to the basic ideas of the SR–based IPMs, in the simplest possible way without unusual, nonstandard notations. On the other hand, there are papers (Terlaky (2001)) and books (Roos et al. (1997)) that discuss the symmetric inequality form of LO problems.

We find the approach for semidefinite optimization, that this discussant has suggested in the discussion paper a very challenging topic for further research. Investigations of SR-IPMs for the suggested classes of optimization problems would not only (re)open the research area that was for the first time discussed in papers Muramatsu and Vanderbei (1993), and Vanderbei and Yang (1995), but also would help to solve the problems in the real world that belong to that class. Due to these two reasons, we are really tempted to work on those problems.

Implementation Issues:

The major issue that most of the reviewers were asking is the computational performance of Self-Regularity based IPMs. Here we briefly discuss some implementation details, for more details readers can consult Salahi et al. (2003) and Zhu et al. (2003). We start by describing some details of the implementation based on the embedding model of the corresponding linear optimization problem.

The implementation of SR–IPMs, like most of the IPMs based packages, is based on the Mehrotra-type predictor-corrector algorithm, and we use the embedding model of the corresponding linear optimization problem.

The following question naturally arises:

How an efficient implementation employs SR search directions?

In what follows we briefly explain the main ideas that guide us when, and how SR search directions are employed in the McIPM package, Zhu et al. (2003).

The algorithm starts with the predictor step, which is the so-called primal–dual affine scaling step. If the maximum step size in this step is sufficiently big, then the algorithm makes Mehrotra's corrector step by using the classical Newton direction that follows by a backtracking line search technique to keep the iterates in a certain neighborhood of the central path. If the maximum feasible step size in the predictor step is not sufficiently big (the current point is too close to the boundary of the feasible set), then because of the strong barrier property of the SR functions, it is reasonable to expect that employing SR search directions bring the iterate close to the central path.

The algorithmic idea, that is explored in Zhu et al. (2003), is based on the theoretical results that have been observed in Peng and Terlaky (2002) and Salahi and Terlaky (2004). If the maximum step size in the predictor step is less than a certain threshold, then McIPM employs a SR search directions to bring the iterate closer to the central path. This is done simply by increasing the barrier degree, q, of the corresponding SR function (McIPM is using the $\Gamma_{1,q}$ family) and choosing the μ value that allows the best possible re–centering. Based on the results of Peng and Terlaky (2002) and Salahi and Terlaky (2004), the closeness of a point to the central path is measured by the ratio $\frac{\mu_g}{\mu_h}$ that is always greater than or equal to one. If this ratio is small enough, then the iterate is sufficiently centered, then we move to the next step, otherwise we have to increase the barrier degree q. The target value in this centering step is chosen adaptively, it is equal to μ_q^*. Since in this case the algorithm tries to bring the iterate close to the central path, by choosing μ_q^* as the target value implies that the duality gap is not changing in this step (see Corollary 2.10). Because the centering step is not expected to be close to the boundary, therefore the algorithm employs a forward tracking inexact line search instead of the backward tracking line search strategy used in the Mehrotra step. If the SR step still results in a too short step, then the value of q is increased until it reaches its maximal allowed value that, due the numerical difficulties, is set to 5.

The computation of the search directions in the McIPM package is based on the normal equation approach and it uses sparse Cholesky factorization. In solving the normal equation, there are several techniques that help us to get an accurate and stable solution quickly. For example, one need to identify and separate dense and sparse columns. Dense columns are separated and dealt with them by using low rank updates (Andersen et al. (2004)). For full details regarding the implementation of the SR–based McIPM package the reader is referred to Zhu et al. (2003).

In Zhu et al. (2003) all the NETLIB test sets for linear optimization problems are solved. Here we give some illustrative numerical results for a selection of large scale problems. Table R.1 (borrowed from Zhu et al. (2003)) shows the superiority of the dynamic use of the SR search directions compared with the fixed value of q during the algorithm.

We have to mention that the implementation of SR–IPMs is also extended to *Quadratic and Second Order Conic* optimization problems. Improvements, due to the use of the SR search directions is also observed for these classes of optimization problems. For details and numerical results the reader is referred to the M.Sc. thesis Romanko (2004) and Wang (2003).

Finally, an infeasible variant of SR-based IPM is also implemented and compared with some well know packages. The results show that the new SR–based implementation is competitive with those packages. The interested reader can consult Salahi et al. (2003) for both the theoretical and computational details of the infeasible variant's implementation.

References

Andersen E.D., Gondzio J., Mészáros Cs. and Xu X. (1996). Implementation of interior point methods for large scale linear programming, In: Terlaky T. (eds.), *Interior Point Methods of Mathematical Programming*. Kluwer Academic Publishers, 189–252.

Andersen E.D., Roos C., Terlaky T., Trafalis T. and Warners J.P. (2004). The use of low-rank updates in interior-point methods. In: Yuan Y. (ed.), *Numerical Linear Algebra and Optimization*, Science Press, 1–12.

Castro J. (2000). A specialized interior point algorithm for multicommodity network flows. *SIAM Journal on Optimization* 10, 852–877.

Gondzio J. (1996). Multiple centrality corrections in a primal–dual method for linear programming. *Computational Optimization and Applications* 6, 137–156.

Gupta A. (2000). WSMP: Watson Sparse Matrix Package (Part I: direct solution of symmetric sparse systems). Technial Report RC 21886 (98462), IBM T.J. Watson Research Center, Yorktown Heights, NY. `http://www.cs.umn.edu/~agupta/wsmp.html`.

Mehrotra S. (1992). On the implementation of a primal-dual interior point method. *SIAM Journal on Optimization* 2, 575–601.

Muramatsu M. and Vanderbei R.J. (1993). Primal–dual affine–scaling algorithms fail for semidefinite programming. *Mathematics of Operations Research* 24, 149–175.

Peng J. and Terlaky T. (2002). A dynamic large-update primal-dual interior-point method for linear optimization. *Optimization Methods and Software* 17, 1077–1104.

Portugal L., Bastos F., Júdice J., Paixão J., Terlaky T. (1996). An investigation of interior-point algorithms for the linear transportation problem. *SIAM Journal on Computing* 17, 1202–1223.

Romanko O. (2004). An Interior Point Approach to Quadratic and Parametric Quadratic Optimization. M.Sc. Thesis, Advanced Optimization Lab, Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada. `http://www.cas.mcmaster.ca/$\sim$oplab/publication`.

Roos C., Terlaky T. and Vial J.-Ph. (1997). *Theory and Algorithms for Linear Optimization: An Interior-Point Approach,* John Wiley.

Salahi M., Terlaky T. and Zhang G. (2003). The complexity of Self-Regular proximity based infeasible IPMs. Technical Report 2003/3, Advanced Optimization Lab. Department of Computing and Software, McMaster University. `http://www.cas.mcmaster.ca/$\sim$oplab/publication`. Submitted to *Computational Optimization and Applications*, 2003.

Salahi M. and Terlaky T. (2004). An adaptive self-regular proximity based large-update interior point methods for linear optimization. *Optimization Methods and Software* (to appear).

Terlaky T. (2001). An easy way to teach interior point methods. *European Journal of Operartional Research* 130, 1–19.

Vanderbei R.J. and Yang B. (1995). The simplest semidefinite programs are trivial. *Mathematics of Operations Research* 20, 590–596.

Wang B. (2003). Implementation of Interior-Point Methods for Second-Order Conic Optimization. M.Sc. Thesis, Advanced Optimization Lab, Depart-

ment of Computing and Software, McMaster University, Hamilton, Ontario, Canada.http://www.cas.mcmaster.ca/$\sim$oplab/publication.

Zhu X., Peng J., Terlaky T. and Zhang G. (2003). On implementing self-regular proximity based feasible IPMs. Technical Report 2003/2, Advanced Optimization Lab. Department of Computing and Software, McMaster University. http://www.cas.mcmaster.ca/$\sim$oplab/publication. Submitted to *INFORMS Journal on Computing*.

Problem	$q=1$		$q=2$		$q=3$		$q=4$		Dynamic		
	Iter	dig.	Iter	dig.	Iter	dig.	Iter	dig.	Iter	dig.	#SR
80bau3b	41	8	42	8	48	8	52	8	40	8	6
bnl1	32	7	33	7	36	7	35	7	33	7	8
cycle	39	7	36	8	38	8	39	8	38	7	8
czprob	36	8	31	9	32	9	32	9	34	9	19
d2q06c	43	7	43	7	48	7	55	7	42	7	17
dfl001	46	6	25	1	25	1	47	6	46	6	6
finnis	26	7	27	7	26	7	29	7	25	7	3
greenbea	48	3	52	3	56	3	61	3	47	3	24
maros	31	5	34	5	37	5	44	5	30	4	3
perold	43	7	44	7	45	7	46	7	44	7	12
pilot	75	7	55	7	53	6	52	6	55	7	24
pilotwe	42	6	42	6	41	6	46	6	40	6	2
pilot4	36	6	35	6	39	6	53	6	34	6	4
pilot87	73	5	77	5	75	5	86	5	71	5	38
scfxm2	27	7	27	7	28	7	30	7	26	7	3
shell	24	9	25	9	32	8	28	8	23	9	2
ship04s	17	8	17	10	17	9	19	8	16	8	1
ship08l	19	9	20	10	20	10	22	9	19	9	2
ship08s	17	9	17	9	19	9	20	9	17	9	1
ship12l	28	9	28	9	28	9	31	8	27	8	10
stocfor2	31	7	31	7	30	7	39	7	28	4	6
cplex2	39	0	37	0	36	0	54	0	34	0	14
refinery	14	0	14	0	13	0	17	0	13	0	1
osa-14	40	8	45	8	47	8	45	9	38	8	17
pds-20	79	8	81	9	81	9	100	8	77	8	64
baxter	52	7	51	7	55	8	57	7	49	7	29

Table R.1: *The Performance of different choices of SR-proximity function*